

Comparative Study Based on De-Occlusion and Reconstruction of Face Images in Degraded Conditions



Laila Ouannes, Anouar Ben Khalifa*, Najoua Essoukri Ben Amara

Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS-Laboratory of Advanced Technology and Intelligent Systems, Sousse 4023, Tunisia

Corresponding Author Email: anouar.benkhalifa@eniso.rnu.tn

<https://doi.org/10.18280/ts.380305>

ABSTRACT

Received: 16 December 2020

Accepted: 5 June 2021

Keywords:

face recognition, degraded conditions, face detection, face de-occlusion, face reconstruction, Laplacian pyramid blending, CycleGANs

In the recent years, the face recognition task has attracted the attention of researchers due to its efficiency in several domains such as surveillance and access control. Unfortunately, there are multiple challenges that decrease the performance of face recognition. Partial occlusion is the most challenging one since it often causes a great lack of information. The main purpose of this paper is to prove that facial reconstruction improves the results of facial recognition compared to de-occlusion and full-face recognition in the presence of occlusion. Our objective is to achieve occluded-face recognition, de-occluded-face recognition, and reconstructed-face recognition. Regarding face reconstruction, we introduce two different methods based on Laplacian pyramid blending and CycleGANs. In order to validate our work, we perform two different feature extraction techniques: hand-crafted features and learned features exploiting the final layers of a pre-trained deep architecture model. The experimental results on the EURECOM Kinect Face Dataset (EKFD) and the IST-EURECOM Light Field Face Database (IST-EURECOM LFFD) show that the proposed face reconstruction approach, compared with the face de-occlusion and occluded-face recognition ones, clearly improves the face recognition task. Our method boosts the classification performance in comparison with the state-of-the-art methods, achieving 94.66% on EKFD and 72.35% on IST-EURECOM LFFD.

1. INTRODUCTION

With the large spread of coronavirus cases, wearing masks has become compulsory in all public spaces, and all breaches of this law will be penalized. This obligation causes the partial occlusion of a significant part of the face, which inhibits face recognition for video surveillance and access control [1, 2].

Compared to several biometric recognition modalities El Kissi Ghalleb and Amara [3]; Mhenni et al. [4]; Rzouga Haddada and Essoukri Ben Amara [5] that expose the weaknesses for authentication, face modality Ouannes et al. [6] is the main natural biometric measure that makes it efficient with a high level of security. Therefore, over the last few years, facial recognition systems have gained wide attention from many corporations and government organizations, notably for surveillance purposes [7].

In earlier work, most researchers focused on facial recognition under controlled conditions such as the frontal pose and the close distance to the camera. However, these researchers no longer meet the requirements for facial recognition in degraded conditions [8]. In fact, four major challenges affect the robustness and performance of facial recognition. The first one is facial expression variations which affect the performance of facial recognition systems, because the appearance of faces is directly affected by facial expressions [9]. The second challenge is illumination variations, which can be caused by either lighting, camera characteristics or skin reflection properties [10]. The third one is head pose variations, which introduce self-occlusion and

projective deformations [11]. Finally, partial occlusion can be engendered by an object such as a cap or scarf, or even a face can be occluded by another face in a crowd of people [12]. Such challenges cause many degradations and lack of information, which leads to a fall in the recognition rate. More details on face recognition in degraded conditions were explained in some previous work [6, 8].

To overcome these problems, several approaches have been proposed. These approaches can be clustered into three main families: global approaches such as PCA, LDA and ICA, local approaches like LBP, the Histogram of Oriented Gradient (HOG) and the Speeded-Up Robust Feature (SURF), and hybrid approaches which combine both global and local approaches to improve the performance of face recognition [13]. All these methods are still valid for some variations but not robust for others, especially in the presence of partial occlusion [14]. Accordingly, recent studies have put forward face de-occlusion [15] and face completion [16]. As a matter of fact, face de-occlusion consists in removing occluded face parts and proceeding with the recognition task only with non-occluded parts. This solution needs the detection and localization of occluded parts [17]. Some recent researchers [16] have conceived the former solution as a loss of information. Hence, they have chosen to head towards face completion, which involves eliminating occluded parts and substitute them with corresponding non-occluded parts from neutral faces. Indeed, there are a lot of techniques for face completion or reconstruction, such as inpainting [18], Laplacian pyramid blending [19], spectral-graph-based

algorithms [20], and robust PCA [21]. In the same vein, with the fast development of deep learning architectures in recent years, several researchers have used Long Short-Term Memory (LSTM) [22], Generative Adversarial Networks (GANs) [23], etc.

In this work, we propose a framework based on two principal steps: face detection and face recognition. The face recognition step is clustered into three main blocks: occluded-face recognition, de-occluded-face recognition, and reconstructed-face recognition. The first step is to detect the face in the image. For this, the Viola & Jones detector Besnassi et al. [24] is combined with the human skin color detection based on the HSV color space [25]. This combination reinforces face detection against face degradation, such as partial occlusion and head pose variations. For face reconstruction, we opt for two different methods: Laplacian pyramid blending Burt and Adelson [19] and Cycle GANs (CycleGANs) [26]. For the three ways of face recognition, different descriptors and classifiers are used in order to validate our approach. For feature extraction, we utilize two different features: (1) hand-crafted features by exploiting the most used and efficient descriptors, which are the SURF Özdemir et al. [27] and the HOG [28], and (2) learned features extracted from the final layers of the pre-trained Inception-v3 architecture model [29]. As regards classification, our suggested method is evaluated by using the k Nearest Neighbors (k-NN) Cunningham and Delany [30] combined with the K Dimensional Tree (KD-Tree) [31] and Support Vector Machine (SVM) [32] classifiers as classical methods and the pre-trained Inception-v3 model as a deep learning method.

The main contribution of our work can be explained as follows:

- First, a comprehensive review is provided, which enumerates the various studies in the recent literature based on occluded-face recognition, face de-occlusion, and face reconstruction.
- Second, we propose a new face detector which is a combination of the Viola and Jones detector with the HSV color space human skin detector, to enhance face degradations and essentially partial occlusion.
- Third, we put forward a comparative study of occluded faces, de-occluded faces and reconstructed faces for the recognition task.
- Finally, to check that face reconstruction improves person authentication, we use several techniques and approaches and we evaluate them on two public datasets.

The rest of this paper is organized as follows: The related work is presented in section 2, in which the state of the art of occluded-face recognition and existing methods for face de-occlusion and face reconstruction is introduced. In section 3, we describe the different methods and techniques used to validate our approach. The evaluation of the different results on the EURECOM Kinect Face Dataset (EKFD) [33] and the IST-EURECOM Light Field Face Database (IST-EURECOM LFFD) [34] is shown in section 4. The conclusion is the subject of section 5.

2. RELATED WORK

Actually, face recognition is a challenging task due to the

high number of present issues, particularly occlusion. For this reason, partial occlusion has attracted the attention of researchers for decades [35]. Advantageously, the more the technology evolves, the more effective solutions to this problem. In this section, different solutions in the literature have been elaborated to overcome the occlusion problem. Three main axes are considered in our proposed approach for facial recognition, which are: occluded faces, de-occluded faces, and reconstructed faces. Therefore, we have divided this section into these same axes.

2.1 Occluded-face recognition

Initially, researchers targeted the recognition of faces with the presence of occlusion [13]. Zhang and Wang [28] proposed a face recognition system based on a HOG-LBP and sparse representation. First, they extracted the HOG and LBP features in order to get the HOG-LBP joint feature that would be projected onto the feature subspace using PCA. This projection was performed to develop an occlusion dictionary. Finally, they exploited sparse representation for the classification task. For catching the contextual and local information for the recognition of occluded faces, Zheng et al. [36] put forward a feature extraction approach. Indeed, they suggested a fusion method to integrate manifold features, namely connected-granule labeling features, structural element features, and reinforced centrosymmetric local binary patterns. An occlusion-aware face recognition approach was propounded by Xu et al. [37] so as to discriminatively learn facial templates in spite of partial occlusion. For this purpose, they introduced an attention mechanism that extracted a local identity-related region. Then, by using global representation, they regrouped local features in order to shape a single template. Another solution to recognize occluded faces was the pairwise differential Siamese network proposed by Song et al. [38], which put up the similarity between occluded facial blocks and falsified feature elements.

All these methods have been quite satisfying for some challenges like head pose variations. However, they deserve to be improved especially in the presence of partial occlusion.

Thus, as research is progressing, researchers choose to exploit only the non-occluded part of the face, believing that face de-occlusion increases the recognition performance.

2.2 De-occluded-face recognition

A Variable-threshold Robust PCA (VRPCA) method was introduced by Leow et al. [21] for eliminating partial occlusion from faces. Indeed, this method could guard the non-occluded parts of the target image with a pretty zero error. Two stage occlusion-aware GANs were suggested, where the former was dedicated to removing partial occlusion [15]. It would be then the supplementary input of the latter GANs for synthesizing the ultimate de-occluded face. To reinstate partially occluded faces, Zhao et al. [22] proposed the LSTM-autoencoder model, which consisted of two LSTM components: The first one aimed to encode occlusion-robust faces, while the second one removed the recurrent occlusion. The deep convolutional generative adversarial networks were introduced by Xu et al. [39] for eliminating partial occlusion from face images and filling up the occluded parts concurrently in an iterative way.

The de-occlusion techniques improved face recognition performances compared with occluded-face recognition, but these results were non-satisfying and had to be ameliorated.

In fact, the fall in the recognition rate was due to the fact that identification was realized by using only the non-occluded part face and not the whole face. In addition, the remaining non-occluded part depended on the location of the occlusion and its size, which influenced the recognition performance.

2.3 Reconstructed-face recognition

To better improve the face recognition results, a lot of researchers have focused on face reconstruction. In order to achieve this task, there have been several methods and techniques such as inpainting [18] and face completion [16]. Afifi et al. [40] proposed, for image blending, a modified Poisson blending technique as a gradient domain. Indeed, for minimizing blending artifacts, the boundary pixels of the source image and the target image were used in the blending process. To inpaint and recognize occluded faces simultaneously, identity preserving generative adversarial networks were suggested by Li et al. [41]. They incorporated an inpainting network, a global-local discriminative network, a parsing network, and an identity network. The main role of the propounded network was to delete the identity diffusion between the real face and its inpainted corresponding. An alternative solution for face completion was Laplacian pyramid blending. After face detection and tracking, if the user chose face replacement, Cao and Liu [42] would blend the face by performing modified Laplacian pyramid blending in order to decrease the variance of face skin colors. To synthesize

realistic examples of face images, Banerjee et al. [43] used Laplacian pyramid blending. Before this step, the authors performed triangulation and then combined these triangles to obtain synthetic faces. The Laplacian pyramid based convolutional network was put forward to complete missing regions by predicting them under multiple resolutions [44]. The more advanced the research, the more effective the solutions. Hence, models and deep architectures were more and more efficient, especially when it came to the reconstruction of occluded faces. Recently, some researchers have considered GANs, to improve their performance. They put forward modifications, which created multiple variants of this model [23]. Ge et al. [45] put forward identity-diversity inpainting to facilitate face recognition under occlusion. Indeed, the authors merged GANs with a pre-trained CNN recognizer. Thus, the proposed approach guaranteed face reconstruction from GANs and its representation from CNNs, which would ensure both challenging tasks: face inpainting and recognition. For key-point guided image generation, a cycle-to-cycle GAN was introduced by Tang et al. [46]. The interactive manner exploration of the joint exploitation of the key-points and the image data would make the proposed network a cross-modal framework. For simultaneously synthesizing and recognizing faces under large-pose variations and partial occlusion, Duan and Zhang [47] propounded a boosting GAN network. Its main aspects were knowledge boosting and an aggregation structure integrated with a boosting network.

Table 1. Summary table for different facial recognition approaches and a selection of related work

Approaches	Related work	Techniques	Datasets	Recognition rate (%)
Occluded-face recognition	Amanzadeh et al. [35]	Improved KED	CAS-PEAL	94.41
	Ouannes et al. [8]	SURF + k-NN & KD-Tree	IST-EURECOM LFFD	80.40
	Xu et al. [37]	OREO	UHDB-31	95.00
			IJB-C	97.74
	Zhang and Wang [28]	HOG-LPB	ORL	88.50
			AR	Sunglasses 98.30 Scarf 98.00
	Song et al. [38]	Pairwise differential Siamese network	LFW	99..20
			MegaFace Challenge	74.40
	AR	Sunglasses 98.19 Scarf 98.33		
	Wan et al. [48]	MaskNet	AR	93.80
Min et al. [33]	LGBP	EKFD	95.05	
De-occluded-face recognition	Cai et al. [49]	OA-GAN	CelebA	83.00
	Dong et al. [15]	OA-GAN	CK+	-
			CelebA	-
	Xu et al. [39]	DCGANs	AR	-
	Zhao et al. [22]	RLA	LFW	-
Leow et al. [21]	VRPCA	CMU-Multi-PIE	-	
Reconstructed-face recognition	Cai et al. [50]	FCSR-GAN	CelebA Helen	-
	Ge et al. [45]	ID-GAN	LFW	96.53
	Li et al. [41]	IP-GANs	LFW	85.50
	Xiong et al. [51]	DRCBNN	CelebA	-
			UTKFace	-
	Wang et al. [44]	Laplacian pyramid based generative framework	CelebA	-
	Deng et al. [52]	GL	AR	Sunglasses 76.60 Scarf 60.90
	Deng et al. [53]	Guided label-learning	AR	Sunglasses 76.70 Scarf 60.80

Note: (-) in Table 1 means that the recognition rate is not mentioned; instead, the authors use the PSN rate and the SSIM index for face reconstruction quality.

The reconstruction techniques and variants provide a clear improvement compared with de-occlusion and occluded-face recognition. Table 1 provides a summary overview of different techniques used in occluded-face recognition studied above, in addition to some related work exploiting the recognition of de-occluded and reconstructed faces.

In this paper, we focus on the impact of the presence of partial occlusion on the face recognition task. Thus, to be able to decide whether the occlusion affects recognition or not, we consider three different axes. The first axis is to recognize the person in the presence of occlusion. The second one is face de-occlusion, which achieves recognition only using the non-occluded face part. The third axis is the reconstruction of the occluded face from the neutral face. To accomplish this task, two different approaches are considered: Laplacian pyramid blending [19] and a CycleGANs model [26].

3. PROPOSED APPROACH

The main goal of this work consists in studying the impact of occlusion on facial recognition. For this reason, we evaluate face recognition with occlusion, after de-occlusion, and after

reconstruction. Figure 1 presents the flowchart of our work. We have to mention that for face reconstruction using CycleGANs, the face detection step is performed after the reconstruction process. This choice is adopted to ensure better quality of images reconstruction. In section 3.1, our approach for face detection is explained. In section 3.2, we introduce the de-occlusion task. Face reconstruction is described in section 3.3, first using the Laplacian pyramid blending technique and second utilizing the CycleGANs deep learning architecture. In section 3.4, we explain the feature extraction and the classification.

3.1 Face detection

In this section, our face detection model is explained. Indeed, face detection is a primordial step for face recognition. The Viola & Jones (VJ) face detector is the most used and effective one for face detection [24]. As for the effectiveness of the VJ detector, it refers to the use of the integral image and Haar-like features. Indeed, the VJ algorithm has four main stages: selecting Haar features, creating integral images, AdaBoost training, and cascading classifiers.

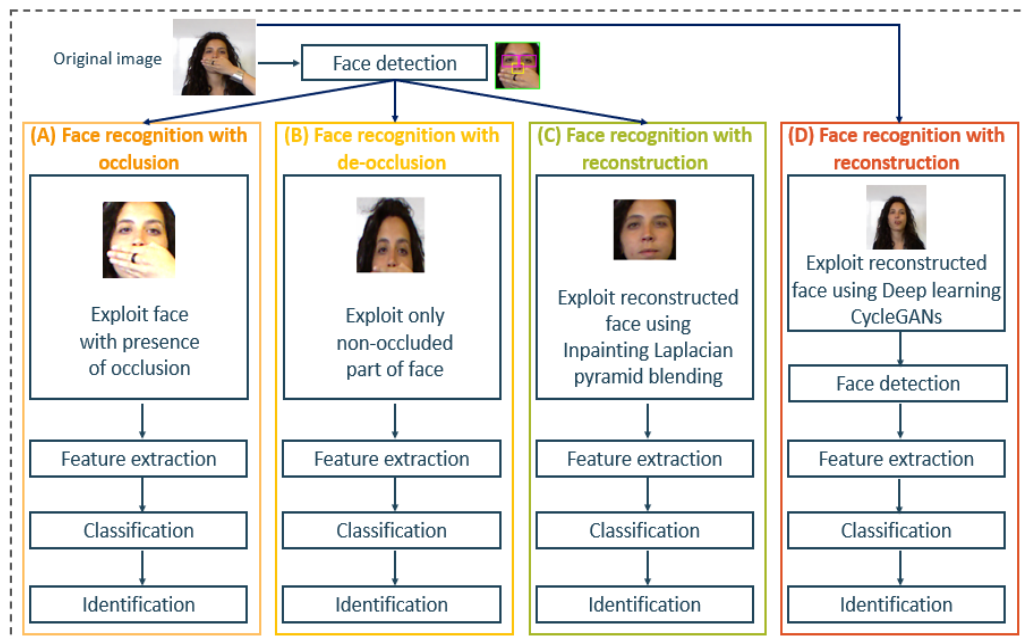


Figure 1. Proposed facial recognition flowchart. After face detection, face recognition is carried out in three different ways: Face recognition with (A) occlusion, (B) de-occlusion, and (C) with reconstruction. For face recognition with reconstruction using CycleGANs (D), face detection is performed after the reconstruction step. were knowledge boosting and an aggregation structure integrated with a boosting network

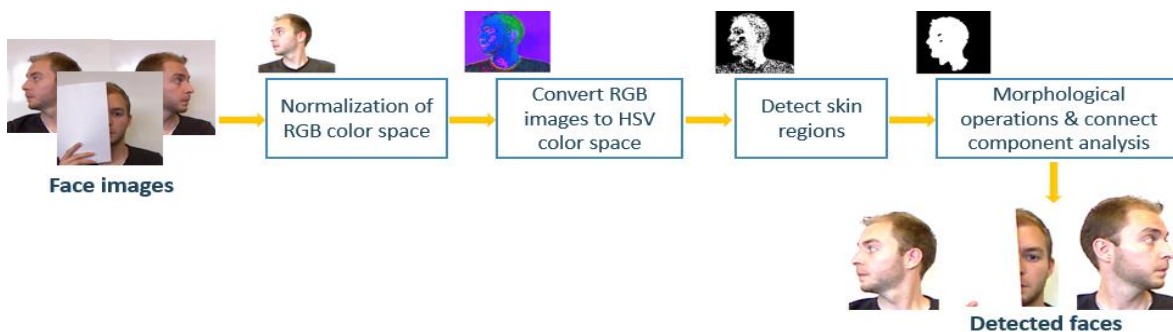


Figure 2. HSV color space steps for face detection based on human skin color detection

Although the VJ framework is very efficient for face detection in controlled conditions, it is still unable to detect faces with some occlusion and orientations. Therefore, due to these limits, we choose to reinforce our face detector by using one of the most used human skin color detection. Indeed, there are many color spaces for skin color detection such as the RGB based color space (RGB, normalized RGB), the hue-based color space (HSV, HSI, and HSL), the luminance-based color space (YCbCr, YUV, and YIQ), and the perceptually uniform color space (CIEXYZ, CIELAB, and CIELUV) [54]. The RGB color space is not preferred for color-based detection because of chrominance, luminance information, and its non-uniform characteristics. Despite the easy separation of the color information from the intensity information in a normalized RGB color space, under unequal illumination conditions, it is not considered for color skin detection. Hence, because of variable lighting and contrast, the RGB color space cannot be used for skin detection under such conditions. Thus, the RGB input image must be converted to another color space that is invariant or at least insensitive to lighting changes. The transformation of the RGB color space to the HSV color space is invariant to ambient light and surface orientations relative to the light source [55]. Luminance and hue-based approaches discriminate color and intensity information even under uneven illumination conditions. Indeed, *hue* points out the dominant color of an area, *saturation* calculates its colorfulness in proportion to its brightness, and *value* denotes the color luminance. Moreover, the separation between luminance and chrominance makes this color space popular in skin color detection. For that, the HSV color space can be a good choice for the human skin detection technique [25]. The transformation of the RGB color space into an HSV color space is assured by using Eq. (1), Eq. (2), and Eq. (3).

$$H = \frac{\frac{1}{2}(2R - G - B)}{\sqrt{(R - G)^2(R - B)(G - B)}} \quad (1)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (2)$$

$$V = \max(R, G, B) \quad (3)$$

The process to ensure the human skin color detection HSV color space is described in Figure 2. First, a normalization of the RGB color space is carried out. Then, the normalized RGB images to the HSV color space are transformed using H, S and V respectively in Eq. (1), Eq. (2) and Eq. (3). The next step is the skin region detection. Indeed, the adequate values of H and S to detect the skin regions are given by Eq. (4) and Eq. (5).

$$0.01 \leq H \leq 0.1 \quad (4)$$

$$0.1 \leq S \leq 0.9 \quad (5)$$

By detecting the human skin region, we get a binary image with holes that obstruct the face detection. In order to refine the detected skin regions, the application of morphological operations is recommended [56]. First, a simple hole filling operation is applied. Then, we apply a dilate function to maintain a particular size of the detected individual blobs of the skin. The next step to detect the face region is to label connected components that form the detected skin region.

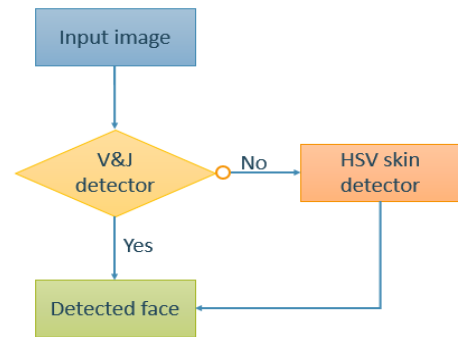


Figure 3. Proposed face detector

The final step is to measure the skin region properties. For that, the area and the bounding box properties are extracted, which ensures skin face detection.

Consequently, our face detection diagram is given by Figure 3. Indeed, if the VJ detector is not able to detect the face region in the image because of the presence of a variation in the pose or partial occlusion, the HSV skin color detector ensures face detection. To the best of our knowledge, this is the first time that these two face detectors are combined, which creates a performed face detector that improves face recognition.

3.2 Face de-occlusion

In this section, we introduce face de-occlusion. We aim to omit the occluded face part and recognize the person only by exploiting the non-occluded face part. To the best of our knowledge, it is the first time that only the non-occluded parts for face recognition are exploited. As described in Figure 4, after face detection, the different faces of the dataset are split into two parts, occluded and non-occluded, which allows our system to detect the occluded part and omit it. The recognition of occluded parts is done by performing feature extraction using the SURF descriptor followed by a classification using the k-NN classifier combined with the KD-Tree. Indeed, this task is like reconstructing a new dataset with no occlusion.

Algorithm 1 describes the de-occlusion system. In fact, face splitting is ensured by exploiting the distances between the eyes, the nose, and the mouth. These distances are provided from the bounding boxes of the VJ detector.

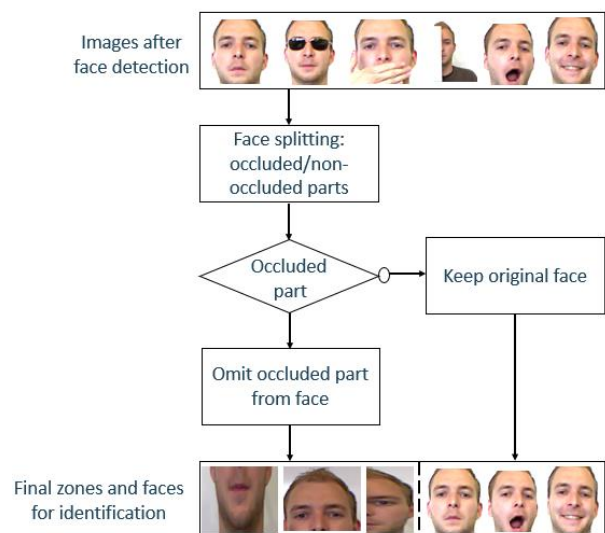


Figure 4. Process of de-occlusion system

Algorithm 1. Face de-occlusion system

```

Require: face with occlusion
Ensure: face without occlusion
for  $i \leftarrow 1$  to width (face) do
  for  $j \leftarrow 1$  to height (face) do
    Split the face into two parts
    if part is non-occluded then
      Label_part  $\leftarrow 0$ 
      Keep the face part
    elseif part is occluded then
      Label_part  $\leftarrow 1$ 
      Remove the occluded part from the occluded face
    end if
  end for
end for
return de-occluded face
  
```

3.3 Face reconstruction

In this section, we aim to reconstruct the occluded faces. In fact, this task completes the face image by omitting the occluded part and replacing it by the corresponding non-occluded part from the neutral face. For face reconstruction [57], two techniques are introduced: Laplacian pyramid blending explained in section 3.3.1 and CycleGANs explained in section 3.3.2.

3.3.1 Laplacian pyramid blending

Before the step of inpainting using the technique of Laplacian pyramid blending, there are many other steps. First, after face detection, we detect whether the face is occluded or not. Hence, if the face is occluded, the occluded part is removed and replaced by the corresponding non-occluded part from the neutral face to get a reconstructed face without occlusion. Reconstruction is performed by using the neutral image as a reference, as described in Figure 5.

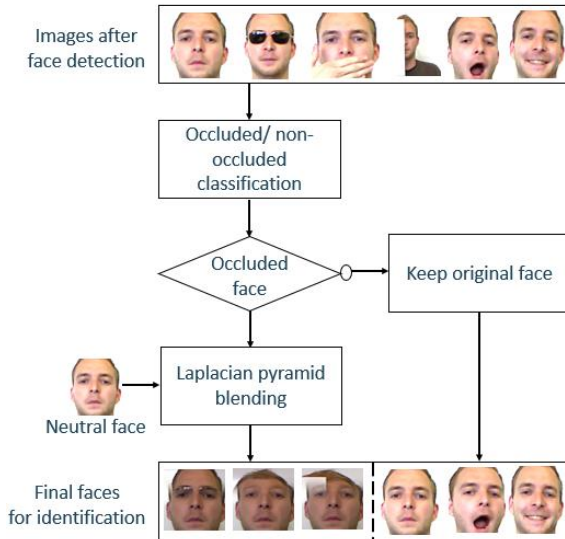


Figure 5. Laplacian pyramid blending reconstruction

In fact, the Laplacian pyramid is built from the Gaussian pyramid. The Gaussian pyramid serves to represent an image at different scales, where the information from the original image is preserved for each scale. In other words, the Gaussian pyramid is a sequence of images, starting with the original, the original narrow by 1/2, the original narrow by 1/4, and so on. At every transition of the pyramid, the image is down-scaled

by a factor of 1/2. By reducing the scale of the image by 1/2, the Gaussian pyramids combine smoothing, using the Gaussian filter, with down-sampling. The convolution of the image's Gaussian smoothing with a Gaussian filter performs a low-pass filter over the image which retains the low-frequency information within an image, while reducing high-frequency information. Simply, using the Gaussian pyramid, the image gets smaller and blurrier at every scale because its size downturns by a factor of 2, and the scale of the applied Gaussian smoothing filter rises by a factor of 2. The next step is Laplacian pyramid construction. At a given level k , the down-scaled image of the next smallest level from the Gaussian pyramid is up-sampled. Then it will be subtracted from the image in the Gaussian pyramid at the current level. The images will have the same size after up-sampling, and the result is the Laplacian image at the current scale. As a matter of fact, Laplacian is in essence a high pass filter. It captures only the details and edges of the images. The construction continues so with the rest of the Laplacian pyramid [19] using Eq. (6).

$$LaplacianImage_k = OriginImage_k - UpsampledGaussian_{k-1} \quad (6)$$

Subsequently, for both input images, the occluded face and the neutral face, a Laplacian pyramid must be built. Then, we construct a Gaussian pyramid for the mask. The application of the Gaussian filter to the subsampled mask makes the blended image smooth. The mask serves to help us combine the Laplacian pyramids for the two inputs. Using an $\alpha + (1 - \alpha)$ combination at each scale, Eq. (7) describes the combined Laplacian at scale k [19].

$$CombinedLaplacian_k = LaplacianA_k * mask_k + LaplacianB_k * (1 - mask_k) \quad (7)$$

Finally, the original image at each scale is reconstructed by adding the combined Laplacian image to the original Gaussian resized image multiplied by their respective masks.

This operation is executed repeatedly, up-sampling the result and adding it to the combined Laplacian, until having the blended image at the original scale. Algorithm 2 explains the face reconstruction system using Laplacian pyramid blending.

Algorithm 2. Face reconstruction system based on Laplacian pyramid blending

```

Require: face with occlusion
Ensure: ReconstructedFace
if face is non-occluded then
  Label_face  $\leftarrow 0$ 
  Keep the face without reconstruction
elseif face is occluded then
  Label_face  $\leftarrow 1$ 
  X  $\leftarrow$  neutral face
  Y  $\leftarrow$  occluded face
  Level_pyramid  $\leftarrow 6$ 
  Mask_y  $\leftarrow$  zeros (size(Y))
  Mask_x  $\leftarrow 1 - Mask_y$ 
  Image_Laplacian_y  $\leftarrow$  PyramidGeneration (Y, level)
  Image_Laplacian_x  $\leftarrow$  PyramidGeneration (X, level)
  Mask_y_filter  $\leftarrow$  GaussianFilter (Mask_y)
  Mask_x_filter  $\leftarrow$  GaussianFilter (Mask_x)
  for  $p \leftarrow 1$  to level do
  
```

```

ImagePyramid ← Image_LaplacianY * MaskY_filter
+Image_LaplacianX * MaskX_filter
end for
ReconstructedFace ← PyramidReconst (ImagePyramid)
end if
return ReconstructedFace

```

3.3.2 CycleGANs

The conditional GANs (cGANs) were proposed for the first time by Goodfellow et al. [23]. Actually, cGANs are composed of two neural networks: the generator and the discriminator. The generator aims to generate new data similar to the expected one. However, the discriminator's goal is to identify if an input image is 'real' or 'fake'. In fact, the discriminator model scores how 'real' images look, learning to differentiate between generated and real images. This score provides feedback to the generator on how well it is performing. The two models are trained simultaneously, and the feedback loop between these two models improves the performance of each other. One application of the cGANs is the pix2pix model, where image-to-image translation is performed [58]. In fact, training an image-to-image translation model requires a dataset composed of paired examples, which are input images, and the same images with the desired modification that can be used as an expected output image. Such datasets present a challenge since it is difficult to collect the same images under different conditions and with various scenes. Furthermore, using a pix-to-pix model shows that the output image can be changed compared to the specified image. CycleGANs are an extension of pix2pix, where image-to-image translation is performed bidirectionally. Indeed, the CycleGANs were proposed by Zhu et al. [59] to deal with unpaired image-to-image translation.

The basic architecture of the CycleGANs consists of two mapping functions, G and F , such that $G: X \rightarrow Y$ and $F: Y \rightarrow X$, and two corresponding adversarial discriminators D_x and D_y . In fact, the mapping function G translates X into outputs, which are fed through D_y to check whether they are 'real' or 'fake' according to domain Y . However, the mapping function F translates Y into outputs, which are fed through D_x to check if they are identical to domain X . Indeed, for the CycleGANs, one generator gets additional feedback from the other generator. This feedback guarantees that an image generated by a generator is cycled consistently. As a result, consecutively applying both generators on an input image should yield a similar image. Accordingly, the loss functions used by the CycleGANs model make it powerful. In addition to the generator loss and the discriminator loss, it involves one more type of loss named cyclic-consistency loss. In fact, to regularize the mappings, the CycleGANs add two cycle-consistency losses: the forward cycle-consistency loss $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$ and the backward cycle consistency loss $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$. This behavior can be incentivized using a cycle-consistency loss in Eq. (8) [59]:

$$\begin{aligned}
L_{\text{cyc}}(G, F) = & E_{x \sim P_{data}}(x) [\|F(G(x)) - x\|_1] \\
& + E_{y \sim P_{data}}(y) [\|G(F(y)) - y\|_1]
\end{aligned} \quad (8)$$

If a sample is translated from domain X to domain Y using the mapping function G , then it is mapped back to domain X using the mapping function F . Thus, we investigate how close it is to the original sample. In the same way, it calculates the loss incurred by translating a sample from domain Y to domain X and then back again to domain Y . This cyclic loss should be

minimized. An explanatory description of the CycleGANs architecture is given in Figure 6.

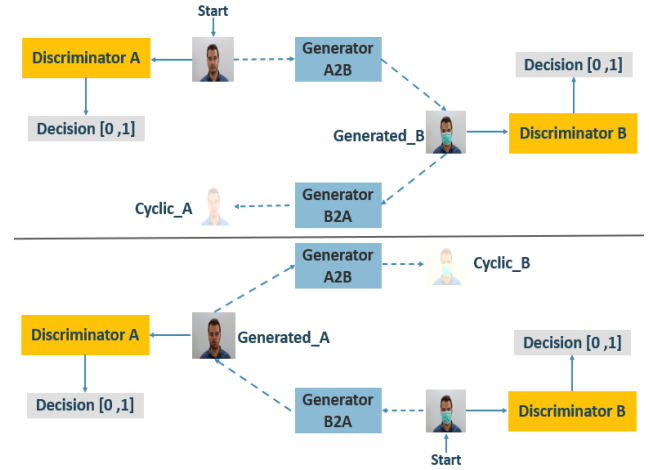


Figure 6. Simplified description of CycleGANs architecture

3.4 Feature extraction and classification

For evaluating our face recognition approaches, we use multiple feature extraction methods and classification techniques. Regarding feature extraction, we use hand-crafted features by exploiting the SURF [60] and HOG [28] descriptors. On the other hand, learned features [61] are extracted from the final layers of the pre-trained Inception-v3 deep architecture model [29]. Actually, with the good reputation of the high-level fine-tuned CNN features, the final convolutional layers of the CNN architecture encode high-level features, allowing these deep layers to be handled for feature extraction [62]. A preprocessing step that converts RGB images to grayscale ones is required before feature extraction. For classification, we use the SVM classifier [32] and the k-NN classifier Cunningham and Delany [30] combined with the KD-Tree [31]. In fact, the KD-Tree is a structure used to organize data in a K-dimensional space according to their spatial distribution. This structure is appropriate to look for closer neighbors, to speed up the search for data in a multi-dimensional space, or to search for intervals. Hence, the exploitation of the KD-Tree facilitates the classification task for the k-NN classifier, which makes the recognition system faster [8]. In addition, as another method for face recognition, the pre-trained Inception-v3 architecture model is exploited. Usually, transfer learning exploits a model that has already been trained on a large dataset for a given classification task. Afterwards, it is used as a new model and trained on a new dataset for the same or another task.

4. EXPERIMENTAL RESULTS

In this section, we extensively evaluate our proposed framework in face recognition over two public datasets. The propounded framework is evaluated from three major aspects: occluded-face recognition, de-occluded-face recognition, and reconstructed-face recognition. All experiments are examined on a 64-bit computer with a CPU intel (R) core (TM) i7-8700 CPU @3.2 GHZ, 16 GB of RAM and an NVIDIA GTX 1660 graphics card with 6GB of VRAM. We build our framework based on the MATLAB R2018b processing environment. Concerning face reconstruction based on CycleGANs models,

the experiments are performed on the Google Collaboratory platform (<https://colab.research.google.com>).

4.1 Datasets

In this section, we evaluate the proposed approach on EKFD [33] and IST-EURECOM LFFD [34]. The former was collected on two different sessions. It consists of 468 images of 52 people (14 females and 38 males) per session. The volunteers were born between 1974 and 1987. They are from various nations with different ethnicities. IST-EURECOM LFFD includes data from 100 volunteers (66 males and 34 females) captured by a Lytro ILLUM camera, with 20 face variations per person, which can be split into four main categories: facial expression variations, head pose variations, illumination variations, and partial occlusion. The volunteers were born between 1957 and 1998, and are from 19 different states. In [34], the authors carefully presented nine facial variations in both sessions: Left Profile (LP), Light On (LO), Neutral (N), Occlusion Eyes (OS), Occlusion Mouth by hand (OH), Occlusion Paper (OP), Open Mouth (OM), Right Profile (RP) and Smile (S).

Figure 7 shows some samples from the two used datasets.

To evaluate our approach, all the images from the first session have been taken to perform the learning phase, whereas the images from the second session have been utilized for the test phase. As an evaluation criterion, we use the facial recognition rate for identification.

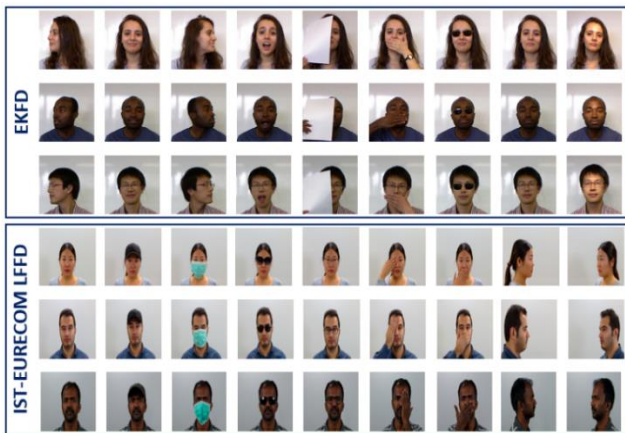


Figure 7. Samples with different skin colors from EKFD and IST-EURECOM LFFD

4.2 Results

As it is known, face detection is a very important step since it influences the performance of face recognition. To make our face detector more efficient, we combine the Viola & Jones detector with the HSV color space human skin detector. The first step of our process is face detection. Figure 8 (B) describes some samples of face detection.

The second step of our process is face recognition. Thus, for face de-occlusion, the face is split into two parts and then these parts are classified into occluded and non-occluded in order to remove the occluded part from the face. Table 2 presents the k-NN and KD-Tree classification rates of the parts using the SURF features.

Afterwards, the type of variations for each person is classified in order to be able to decide whether the face is occluded or not. This classification allows us to know which

face will be reconstructed using the technique of Laplacian pyramid blending. Table 3 provides some experimental results concerning the HOG and SURF descriptors.

It is clear from Table 3 that the SURF descriptor gives better results with the k-NN classifier combined with KD-Tree. Besides the best results shown in Table 3, the choice of the SURF descriptor instead of the HOG descriptor is that SURF allows the use of images with different sizes, which keeps the resolution of images and avoids the waste of information.



Figure 8. Processing of occluded samples of EKFD (the first three left columns) and of IST-EURECOM LFFD (from column 4 to column 9) of different approaches: (A) Original images, (B) Detected faces, (C) De-occluded faces, (D) Reconstructed faces using Laplacian pyramid blending, (E) Reconstructed faces using CycleGANs

Table 2. Classification rates (%) of occluded and non-occluded parts on images of EKFD and IST-EURECOM LFFD

Methods/ Datasets	EKFD	CR (%) IST-EURECOM LFFD
SURF + k-NN & KD-Tree	91.88	88.95

Table 3. Classification rates (%) of different type variations on images of EKFD and IST-EURECOM LFFD

Methods/ Datasets	EKFD	CR (%) IST-EURECOM LFFD
HOG+SVM	82.69	73.50
SURF + k-NN & KD-Tree	92.09	82.35

The different parameters of the various techniques and models are explained as follows. To detect the SURF features, we can switch several parameters such as the strongest feature threshold which should be decreased to return more blobs, the number of scale levels per octave to calculate (ranged from 3 to 6), and the number of octaves to implement (between 1 and 4). For the latter parameter, while varying the number of octaves to increase the number of blobs, the classification results improve further. To extract also the SURF features, we can change the length of the SURF feature vector from 64 to 128 knowing that the larger feature size of 128 provides greater accuracy, but it decreases the feature matching speed. For the HOG descriptor, the size of the HOG cell and the number of cells in the block can be varied. Concerning the k-NN classifier, we can change the number of neighborhoods.

For the KD-Tree, better results can be obtained by changing the maximal number of comparisons. Concerning the SVM classifier, we use a multi class classification with a radial basis kernel function. Coming to the pre-trained Inception-v3 model, it is trained using the Adam optimizer and an initial learning rate equal to 10^{-4} . Then from the final layers of the Inception-v3, we extract learned features. As for Laplacian pyramid blending, the Laplacian pyramid must be constructed. For this, we use six levels and a Gaussian filter with a standard deviation of 2 and a size equal to 9. Moving now to the CycleGANs models, to reconstruct occluded faces, we use the

architecture based on the model introduced by Zhu et al. [59]. The different recognition rate results of the two datasets, EKFD and IST-EURECOM LFFD, are acquired using the SURF descriptor and the k-NN classifier combined with the KD-Tree. Added to that, the recognition rates are received using the pre-trained Inception-v3 model, and several recognition rate results are gained using learned features extracted from the final layers of the pre-trained Inception-v3 model and are classified using the SVM classifier and the k-NN classifier. All these are given respectively in Figure 9 and Figure 10.

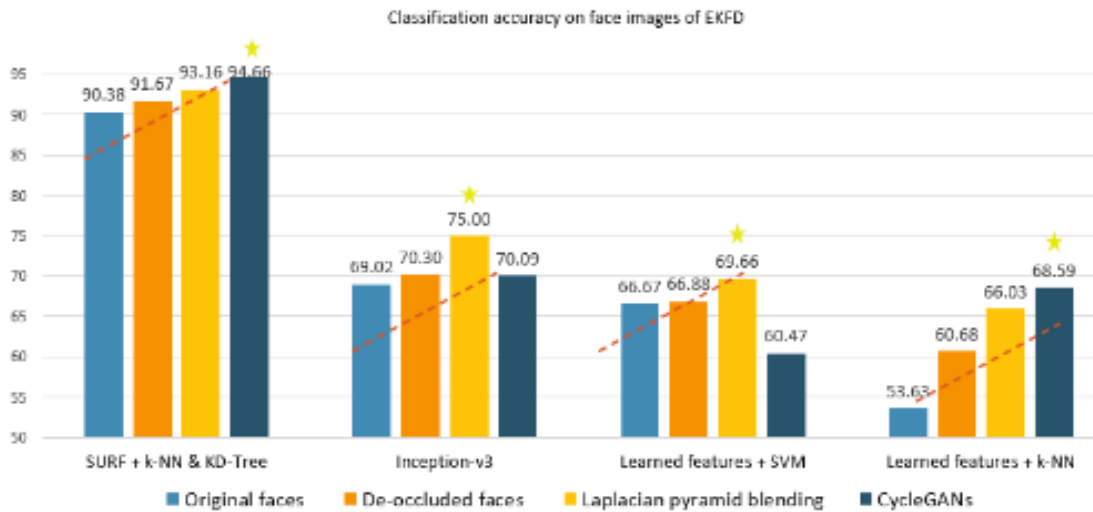


Figure 9. Identification recognition rates (%) on EKFD

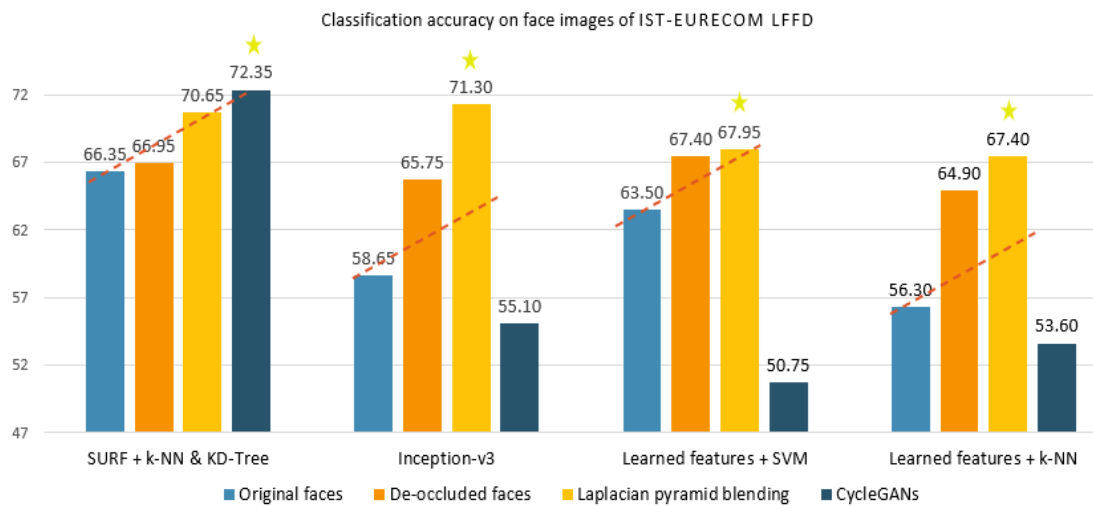


Figure 10. Identification recognition rates (%) on IST-EURECOM LFFD



Figure 11. Reconstructed faces samples using CycleGANs

4.3 Discussion

In this section, five main points can be discussed. Firstly, the recognition results drawn in Figure 9 and Figure 10 show that face reconstruction outperforms other techniques for both used datasets, which confirms that de-occlusion improves the facial recognition rate, but face reconstruction provides better recognition results. Indeed, by using hand-crafted features, face reconstruction based on the CycleGANs method outperforms other approaches by reaching a 94.66% accuracy rate for EKFD and 72.35% for IST-EURECOM LFFD.

Second, when comparing the results given by the pretrained Inception-v3 model and those obtained by learned features with those given by hand-crafted features, a reduction in the results of all approaches is noticed. In fact, this drop can be explained by the small number of samples taken from EKFD and IST-EURECOM LFFD compared to the big number of samples (about thousands and millions) required by deep architectures. Actually, these two datasets are used for the variety of challenges they present, which allows us to evaluate our approaches, especially in the presence of partial occlusion. Nevertheless, the provided results remain quite satisfactory particularly with the presence of occlusion caused by wearing a mask, which has become mandatory due to the Coronavirus pandemic.

Third, the performance of facial reconstruction using CycleGANs presents less efficient results than other techniques. Precisely, these results can be explained by the quality of the reconstruction which sometimes introduces significant changes from the neutral face characteristics used during this step. Some samples are depicted in Figure 11.

Fourth, we have validated our proposed approach on these datasets presenting a large variation of participants' ages, races and skin colors, which proves the robustness of our recognition solution. People with black and yellow skins represent 34% of EKFD and 15% of IST-EURECOM LFFD. Table 4 and Table 5 give the values of the classification rates, which are clustered according the skin colors of the participants from both datasets, respectively.

As noted in Table 4 and Table 5, our model performs better on the black color skin. This is explained by the fact that our system recognition inputs are grayscale images. Obviously, individuals with the black skin are more persistent to conversion from the RGB color space to grayscale one.

Table 4. Classification rates accordingly participants' skin colors from EKFD using SURF + k-NN & KD-Tree

Approaches/ Skin colors	CR (%)		
	White	Yellow	Black
Occluded-face recognition	88.89	91.85	96.30
De-occluded-face recognition	90.19	93.33	96.30
Reconstructed-face recognition	94.77	93.33	100

Table 5. Classification rates accordingly participants' skin colors from IST-EURECOM LFFD using SURF + k-NN & KD-Tree

Approaches/ Skin colors	CR (%)		
	White	Yellow	Black
Occluded-face recognition	64.70	74.28	95.00
De-occluded-face recognition	65.18	76.43	85.00
Reconstructed-face recognition	70.35	82.50	100

Finally, to validate our work on both used datasets, we have compared our work with the existing state-of-the-art methods. To the best of our knowledge, this is the first time that we ensure face recognition from de-occluded faces and reconstructed faces. The recognition rate values are given in both Table 6 and Table 7, where it is clear that our results outperform the state-of-the-art methods and proves that the reconstruction task improves face recognition further.

Otherwise, all our approaches record good rates in both used datasets. In particular, the best results are got using the propounded face reconstruction approach.

Table 6. Summary table for performance evaluation of different methods on EKFD images

Approaches	Methods	Recognition rates (%)
Occluded-face recognition	2D-QSPCA [63]	69.87
	NMR [64]	75.00
	C-means [65]	72.08
	SIFT [33]	74.18
	SURF + k-NN & KD-Tree	90.38
De-occluded-face recognition		91.67
Reconstructed-face recognition		94.66

Table 7. Summary table for performance evaluation of different methods on IST-EURECOM LFFD images

Approaches	Methods	Recognition rates (%)
Occluded-face recognition	PCA [66]	51.50
	LBP [66]	55.60
	SURF + k-NN & KD-Tree	66.35
De-occluded-face recognition		66.95
Reconstructed-face recognition		72.35

5. CONCLUSIONS

In this paper, a comparative study based on the de-occlusion and reconstruction of face images in degraded conditions has been proposed. Indeed, our study consists in inquiring if the occlusion affects the performance of face recognition or not. To the best of our knowledge, it is the first time that such a study is performed. The first step of our pipeline is face detection, which is performed with the suggested face detector: the Viola & Jones detector combined with the HSV human skin detector. On the other hand, the second step is face recognition. For this, we have proceeded in three different approaches. First, we have realized the recognition of faces with occlusion. Second, we have proceeded on de-occlusion, which allows the recognition of only non-occluded face parts. Third, we have carried out face reconstruction using two different methods: Laplacian pyramid blending and CycleGANs. In order to validate our work, several techniques for feature extraction as well as classification have been envisaged. The SURF and HOG descriptors have been used to extract hand-crafted features. The SVM and the k-NN combined with the KD-Tree classifiers have been exploited as classical methods for classification. We have also made use of the pre-trained Inception-v3 deep architecture model for transfer learning and for the extraction of learned features from its final layers. The experimental results demonstrate that occlusion negatively affects face recognition and reconstruction with both Laplacian pyramid blending and CycleGANs techniques, which improves further the face recognition performance. A comparative study on the state-of-the-art methods on the EKFD and IST-EURECOM LFFD datasets shows that our face reconstruction system outperforms other methods.

REFERENCES

- [1] Ben Khalifa, A., Essoukri Ben Amara, N. (2008). Fusion at the feature level for person verification based on off line handwriting and signature. 2008 2nd International Conference on Signals, Circuits and Systems, pp. 1-5. <https://doi.org/10.1109/ICSCS.2008.4746901>
- [2] Ben Khalifa, A., Essoukri Ben Amara, N. (2009). Bimodal biometric verification with different fusion levels. IEEE International Multi-Conference on Systems, Signals & Devices, pp. 1-6. <https://doi.org/10.1109/SSD.2009.4956731>
- [3] El Kissi Ghalleb, A., Essoukri Ben Amara, N. (2017). Remote person authentication in different scenarios based on gait and face in front view. 2017 14th International Multi-Conference on Systems, Signals & Devices (SSD), pp. 486-491. <https://doi.org/10.1109/SSD.2017.8167008>
- [4] Mhenni, A., Cherrier, E., Rosenberger, C., Essoukri Ben Amara, N. (2019). Analysis of Doddington zoo classification for user dependent template update: Application to keystroke dynamics recognition. Future Generation Computer Systems, 97: 210-218. <https://doi.org/10.1016/j.future.2019.02.039>
- [5] Rzouga Haddada, L., Essoukri Ben Amara, N. (2019). Double watermarking-based biometric access control for radio frequency identification card. International Journal of RF and Microwave Computer-Aided Engineering, 29(11): e21905. <https://doi.org/10.1002/mmce.21905>
- [6] Ouannes, L., Ben Khalifa, A., Essoukri Ben Amara, N. (2019). Deep learning vs hand-crafted features for face recognition under uncontrolled conditions. 2019 International Conference on Signal, Control and Communication (SCC), pp. 185-190. <https://doi.org/10.1109/SCC47175.2019.9116159>
- [7] Alphonse, P.J.A., Sriharsha, K.V. (2020). Depth perception in a single RGB camera using body dimensions and centroid property. Traitement du Signal, 37(2): 333-340. <https://doi.org/10.18280/ts.370220>
- [8] Ouannes, L., Ben Khalifa, A., Essoukri Ben Amara, N. (2020). Facial recognition in degraded conditions using local interest points. 17th IEEE International Multi-Conference on Systems, Signals & Devices2020 (SSD2020), pp. 1-6. <https://doi.org/10.1109/SSD49366.2020.9364124>
- [9] Ayeche, F., Alti, A., Boukerram, A. (2020). Improved face and facial expression recognition based on a novel local gradient neighborhood. Journal of Digital Information Management, 18(1): 33. <https://doi.org/10.6025/jdim/2020/18/1/33-42>
- [10] Vishwakarma, V.P., Dalal, S. (2020). A novel non-linear modifier for adaptive illumination normalization for robust face recognition. Multimedia Tools and Applications, 79: 11503-11529. <https://doi.org/10.1007/s11042-019-08537-6>
- [11] Gao, G., Yu, Y., Yang, M., Chang, H., Huang, P., Yue, D. (2020). Cross-resolution face recognition with pose variations via multilayer locality-constrained structural orthogonal Procrustes regression. Information Sciences, 506: 19-36. <https://doi.org/10.1016/j.ins.2019.08.004>
- [12] Zeng, D., Veldhuis, R., Spreuwers, L. (2020). A survey of face recognition techniques under occlusion. arXiv preprint arXiv:2006.11366.
- [13] Mohd Fikri Azli, A., Md Shohel, S., Kalaiarasi Sonai M., Housam Khalifa, B., Afizan, A., Siti Zainab, I. (2014). Face recognition with Symmetric Local Graph Structure (SLGS). Expert Systems with Applications, 41(14): 6131-6137. <https://doi.org/10.1016/j.eswa.2014.04.006>
- [14] Kortli, Y., Jridi, M., Al Falou A., Atri, M. (2020). Face recognition systems: A survey. Sensors, 20(2): 342. <https://doi.org/10.3390/s20020342>
- [15] Dong, J., Zhang, L., Zhang, H., Liu, W. (2020). Occlusion-aware GAN for face de-occlusion in the wild. 2020 IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6. <https://doi.org/10.1109/ICME46284.2020.9102788>
- [16] Mathai, J., Masi, I., AbdAlmageed, W. (2019). Does generative face completion help face recognition? 2019 International Conference on Biometrics (ICB), pp. 1-8. https://doi.org/10.1007/978-3-662-56006-8_2
- [17] Zohra, F.T., Gavrilova, M. (2017). Kinect face recognition using occluded area localization method. Transactions on Computational ScienceXXX. Springer, pp. 12-28. https://doi.org/10.1007/978-3-662-56006-8_2
- [18] Liu, Y., Shu, C. (2015). A comparison of image inpainting techniques. In Sixth International Conference on Graphic and Image Processing (ICGIP 2014), 9443: 94431M. <https://doi.org/10.1117/12.2178820>
- [19] Burt, P., Adelson, E. (1983). The Laplacian pyramid as a compact image code. IEEE Transactions on Communications, 31(4): 532-540. <https://doi.org/10.1109/TCOM.1983.1095851>
- [20] Deng, Y., Dai, Q., Zhang, Z. (2011). Graph Laplace for occluded face completion and recognition. IEEE Transactions on Image Processing, 20(8): 2329-2338. <https://doi.org/10.1109/TIP.2011.2109729>
- [21] Leow, W.K., Li, G., Lai, J., Sim, T., Sharma, V. (2016). Hide and seek: Uncovering facial occlusion with variable-threshold robust PCA. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1-8. <https://doi.org/10.1109/WACV.2016.7477579>
- [22] Zhao, F., Feng, J., Zhao, J., Yang, W., Yan, S. (2018). Robust LSTM-autoencoders for face de-occlusion in the wild. IEEE Transactions on Image Processing, 27(2): 778-790. <https://doi.org/10.1109/TIP.2017.2771408>
- [23] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. (2014). Generative adversarial networks. arXiv preprint arXiv:1406.2661.
- [24] Besnassi, M., Neggaz, N., Benyettou, A. (2020). Face detection based on evolutionary Haar filter. Pattern Analysis and Applications, 23(1): 309-330. <https://doi.org/10.1007/s10044-019-00784-5>
- [25] Dahmani, D., Cheref, M., Larabi, S. (2020). Zero-sum game theory model for segmenting skin regions. Image and Vision Computing, 99: 103925. <https://doi.org/10.1016/j.imavis.2020.103925>
- [26] Zhao, Y., Wu, R., Dong, H. (2020). Unpaired image-to-image translation using adversarial consistency loss. arXiv preprint arXiv:2003.04858,2020.
- [27] Özdemir, H., Sever, R., Polat, Ö. (2019). GA-based optimization of SURF algorithm and realization based on Vivado-HLS. Traitement du Signal, 36(5): 377-382. <https://doi.org/10.18280/ts.360501>
- [28] Zhang, J.X., Wang, J.Y. (2020). Local occluded face recognition based on hog-LBP and sparse representation. 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), pp.

- 808-813.
<https://doi.org/10.1109/ICAICA50127.2020.9182556>
- [29] Szegeedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z. (2016). Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818-2826. <https://doi.org/10.1109/CVPR.2016.308>
- [30] Cunningham, P., Delany, S.J. (2020). K-nearest neighbour classifiers. arXiv preprint arXiv:2004.04523.
- [31] Du, J., Bian, F. (2020). A privacy-preserving and efficient k-nearest neighbor query and classification scheme based on k-dimensional tree for outsourced data. IEEE Access, 8: 69333-69345. <https://doi.org/10.1109/ACCESS.2020.2986245>
- [32] Yu, G., Zhang, Z. (2020). Face and occlusion recognition algorithm based on global and local. Journal of Physics: Conference Series, 1453: 012019.
- [33] Min, R., Kose, N., Dugelay, J. (2014). Kinectfacedb: A kinect database for face recognition. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 44(11): 1534-1548. <https://doi.org/10.1109/TSMC.2014.2331215>
- [34] Sepas-Moghaddam, A., Chiesa, V., Correia, P.L., Pereira, F., Dugelay, J.L. (2017). The IST-EURECOM light field face database. Biometrics and Forensics (IWBF), 2017 5th International Workshop on IEEE, pp. 1-6. <https://doi.org/10.1109/IWBF.2017.7935086>
- [35] Amanzadeh, S., Forghani, Y., Chabok, J.M. (2020). Improvements on learning kernel extended dictionary for face recognition. Revue d'Intelligence Artificielle, 34(4): 387-394. <https://doi.org/10.18280/ria.340402>
- [36] Zheng, W., Gou, C., Wang, F.Y. (2020). A novel approach inspired by optic nerve characteristics for few-shot occluded face recognition. Neurocomputing, 376: 25-41. <https://doi.org/10.1016/j.neucom.2019.09.045>
- [37] Xu, X., Sarafianos, N., Kakadiaris, I.A. (2020). On improving the generalization of face recognition in the presence of occlusions. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 3470-3480. <https://doi.org/10.1109/CVPRW50498.2020.00407>
- [38] Song, L., Gong, D., Li, Z., Liu, C., Liu, W. (2019). Occlusion robust face recognition based on mask learning with pairwise differential Siamese network. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 773-782. <https://doi.org/10.1109/ICCV.2019.00086>
- [39] Xu, L., Zhang, H., Raitoharju, J., Gabbouj, M. (2018). Unsupervised facial image de-occlusion with optimized deep generative models. 2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1-6. <https://doi.org/10.1109/IPTA.2018.8608127>
- [40] Afifi, M., Hussain, K.F. (2015). Mpb: A modified poisson blending technique. Computational Visual Media, 1(4): 331-341. <https://doi.org/10.1007/s41095-015-0027-z>
- [41] Li, C., Ge, S., Hua, Y., Liu, H., Jin, X. (2020). Occluded face recognition by identity-preserving inpainting. Cham: Springer International Publishing, pp. 427-437. https://doi.org/10.1007/978-3-030-04946-1_41
- [42] Cao, Q., Liu, R. (2014). Real-time face tracking and replacement. Google Scholar, pp. 1-10.
- [43] Banerjee, S., Bernhard, J.S., Scheirer, W.J., Bowyer, K. W., Flynn, P.J. (2017). SREFI: Synthesis of realistic example face images. 2017 IEEE International Joint Conference on Biometrics (IJCB), pp. 37-45. <https://doi.org/10.1109/BTAS.2017.8272680>
- [44] Wang, Q., Fan, H., Sun, G., Cong, Y., Tang, Y. (2019). Laplacian pyramid adversarial network for face completion. Pattern Recognition, 88: 493-505. <https://doi.org/10.1016/j.patcog.2018.11.020>
- [45] Ge, S., Li, C., Zhao, S., Zeng, D. (2020). Occluded face recognition in the wild by identity-diversity inpainting. IEEE Transactions on Circuits and Systems for Video Technology, 30(10): 3387-3397. <https://doi.org/10.1109/TCSVT.2020.2967754>
- [46] Tang, H., Liu, H., Sebe, N. (2020). Unified generative adversarial networks for controllable image-to-image translation. IEEE Transactions on Image Processing, 29: 8916-892. <https://doi.org/10.1109/TIP.2020.3021789>
- [47] Duan, Q., Zhang, L. (2020). Look more into occlusion: Realistic face frontalization and recognition with BoostGAN. IEEE Transactions on Neural Networks and Learning Systems, pp. 1-15. <https://doi.org/10.1109/TNNLS.2020.2978127>
- [48] Wan, W., Chen, J. (2017). Occlusion robust face recognition based on mask learning. 2017 IEEE International Conference on Image Processing (ICIP), pp. 3795-3799. <https://doi.org/10.1109/ICIP.2017.8296992>
- [49] Cai, J., Han, H., Cui, J., Chen, J., Liu, L., Zhou, S.K. (2021). Semi-supervised natural face de-occlusion. IEEE Transactions on Information Forensics and Security, 16: 1044-1057. <https://doi.org/10.1109/TIFS.2020.3023793>
- [50] Cai, J., Han, H., Shan, S., Chen, X. (2020). FCSR-GAN: Joint face completion and super-resolution via multi-task learning. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2(2): 109-121. <https://doi.org/10.1109/TBIOM.2019.2951063>
- [51] Xiong, H., Wang, C., Wang, X., Tao, D. (2020). Deep representation calibrated Bayesian neural network for semantically explainable face inpainting and editing. IEEE Access, 8: 13457-13466. <https://doi.org/10.1109/ACCESS.2019.2963675>
- [52] Deng, Y., Dai, Q., Zhang, Z. (2011). Graph Laplace for occluded face completion and recognition. IEEE Transactions on Image Processing, 20(8): 2329-2338. <https://doi.org/10.1109/TIP.2011.2109729>
- [53] Deng, Y., Li, D., Xie, X., Lam, K., Dai, Q. (2009). Partially occluded face completion and recognition. 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 4145-4148. <https://doi.org/10.1109/ICIP.2009.5413474>
- [54] Busin, L., Vandenbroucke, N., Macaire, L. (2008). Color spaces and image segmentation. Advances in Imaging and Electron Physics, 151: 65-168. [https://doi.org/10.1016/S1076-5670\(07\)00402-8](https://doi.org/10.1016/S1076-5670(07)00402-8)
- [55] Hamdini, R., Diffellah, N., Namane, A. (2019). Robust local descriptor for color object recognition. Traitement du Signal, 36(6): 471-482. <https://doi.org/10.18280/ts.360601>
- [56] Thakur, S., Paul, S., Mondal, A., Das, S., Abraham, A. (2011). Face detection using skin tone segmentation. 2011 World Congress on Information and Communication Technologies, pp. 53-60. <https://doi.org/doi:10.1109/WICT.2011.6141217>
- [57] Li, H., Ge, X. (2019). Design and application of an image classification algorithm based on semantic

- discrimination. *Traitement du Signal*, 36(5): 439-444. <https://doi.org/10.18280/ts.360509>
- [58] Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A. (2017). Image-to-image translation with conditional adversarial networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125-1134. <https://doi.org/10.1109/CVPR.2017.632>
- [59] Zhu, J., Park, T., Isola, P., Efros, A.A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2242-2251. <https://doi.org/10.1109/ICCV.2017.244>
- [60] Merino, I., Azpiazu, J., Remazeilles, A., Sierra, B. (2020). 2d image features detector and descriptor selection expert system. arXiv preprint arXiv:2006.02933.
- [61] Tarchoun, B., Ben Khalifa, A., Dhifallah, S., Jegham, I., Mahjoub, M.A. (2020). Hand-crafted features vs deep learning for pedestrian detection in moving camera. *Traitement du Signal*, 37(2): 209-216. <https://doi.org/10.18280/ts.370206>
- [62] Marnissi, M.A., Fradi, H., Dugelay, J. (2019). On the discriminative power of learned vs. hand-crafted features for crowd density analysis. 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, pp. 1-8. <https://doi.org/10.1109/IJCNN.2019.8851764>
- [63] Xiao, X., Zhou, Y. (2019). Two-dimensional quaternion PCA and sparse PCA. *IEEE Transactions on Neural Networks and Learning Systems*, 30(7): 2028-2042. <https://doi.org/10.1109/TNNLS.2018.2872541>
- [64] Yang, J., Luo, L., Qian, J., Tai, Y., Zhang, F., Xu, Y. (2017). Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(1): 156-171. <https://doi.org/10.1109/TPAMI.2016.2535218>
- [65] Oravec, M., Sopiak, D., Jirka, V., Pavlovičová, J., Budiak, M. (2015). Clustering algorithms for face recognition based on client-server architecture. 2015 International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 241-244. <https://doi.org/10.1109/IWSSIP.2015.7314221>
- [66] Sepas-Moghaddam, A., Correia, P.L., Nasrollahi, K., Moeslund, T.B., Pereira, F. (2018). Light field based face recognition via a fused deep representation. 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1-6. <https://doi.org/10.1109/MLSP.2018.85169>