FaRes: A Face Recognition System based on Motion Detection and Image Super Resolution

Lionel Landry Sop Deffo
University of Buea
Buea, Cameroon

Elie Fute Tagne
University of Dschang
University of Buea
Dschang, Buea, Cameroon

ABSTRACT
In Most implementations in recent computer applications include the processing of images and videos. They are therefore subject to many challenges notably the quality of images, the resources availability, variations in scene, etc. With the advent of deep learning approaches, the challenges faced have increased even further: this is because deep learning techniques are mostly based on artificial neural networks and as such, requires a lot of specific resources such as dedicated GPUs (graphical processing units). To cope with these challenges, solutions requiring less resource utilization are proposed. This passes through the inclusion of movement detection module leading to the restriction of further computations to only the zone of interest; that is, the area containing the element of interest (humans, cars, etc.). In addition, image enhancement modules are often used to increase the accuracy of the result. This paper proposes in an approach called FaReS: A face recognition system based on motion detection and image super resolution which uses a proposed improved moving object detection approach, adapted image enhancement and an adapted classification approach. It is assumed that environment considered is that where multimedia sensors such as IP cameras have been installed. The study is focused only on humans and not on all type of objects.

General Terms
Supervised Learning, Background Subtraction, Resources consumption

Keywords
Deep Leaning, Face Recognition, Image Enhancement, Moving object, Multimedia Signal

1. INTRODUCTION
To implement any application on multimedia, signal a certain number of operations need to be performed as far as computer vision is concerned; one of the prior operations is the detection of moving objects in a scene. This usually passes through pixel segmentation enabling the dissociation of the pixels of interest from others. These pixels are usually called foreground pixels, while the others are called background pixels. This has led to the development of many approaches by researchers each of them trying to overcome a challenge in a specific scenario such as; complex background, high illumination, poor quality of input image, etc. The process achieved is called background subtraction or background elimination. The proposed approach will therefore process a collection of images from a video called frames and must be done in enough time for it to be applied in some real time applications.

Among the mentioned operations, there exist detection and localization in a case of specific object processing. If dealing with humans, the detection will include the localization of the human faces in the input frame which processes the whole frame to find the positions of the faces. This raises the problem of resources management because in the case of a multimedia sensors deployed in an environment, the resources available especially memory and computational power are very limited so, certain computations will not be possible. The idea will be to reduce the amount of the multimedia signal to be processed by restricting further computations on the input frame only to the area of interest, and also to improve the quality of the resulting frame in case of low quality, in order to increase the accuracy of the results. This paper presents an approach that achieves these goals and shows the impact of such operations on the resource consumption, as well as the accuracy of the results. It is called FaReS: A **Fa**ce **Re**cognition **S**ystem based on motion detection and image super resolution.

To better present the work, the following structure is used for the remaining aspect of the paper, Section 2 briefly discusses on related work, specifically on background subtraction, face detection and face recognition. Section 3 is dedicated to the presentation of FaReS approach. In Section 4 highlights on different implementations as well as presentation of the results obtained and discussion on their meaning. Conclusion of the work is made in Section 5 that briefly summarizes the findings and open reflections on direction that could be explored in order to improve the work.

2. RELATED WORK
2.1 Background Subtraction algorithms
Background subtraction approaches can be grouped into parametric and non-parametric approaches. While parametric (also called probabilistic approaches) concern approaches that model the background with Normal (Gaussian) distribution over pixel's intensity values of an image [2], non-parametric approaches (usually referred to as sample based approaches) [1] refer to approaches that instead model pixels by samples values and by the mean of specific computations. A pixel is classified as belonging to either background or to the foreground. Some probabilistic approaches include GMM [4], TLGMM [5], STGMM [6] and non-parametric include [1], Bayesian approach [7], the kernel density estimation (KDE) approach [23]. More details on background subtraction approaches can be found in [2], [9].

2.2 Face detection algorithms
As earlier mentioned, moving objects detection is followed by face detection or localization by some authors. These approaches can be classified as Viola John approaches [10], which were one of the first to be proposed. Here, certain patterns called Adaboost[11] are used to verify the presence or the absence of a face in an input frame. Another approach

is skin color approach where [12] to find faces, a method based on color is used in comparing the color of face with that of the background using the fact that human faces have significant color distribution that differs from background objects. Another class is support vector machines approaches [13] with an alternative called appearance approach [14]. Knowledge based approach also exists [12]. More details could be founded in [13], [15].

2.3 Face recognition algorithms

Face recognition approaches are so many, but can be grouped into appearance base approaches, features based approaches and neural network approaches. In appearance-based approaches, the image is represented with a mathematical distribution in order to do a dimensionality reduction; in this view some can been outlined such as Fishers methods [16], Principle Component Analysis [17] and support vector machines [13]. Features-based approach on the other hand, tries to find a vector usually called features vector for each sample in the knowledge base. Some approaches related to that are Hidden Markov Model (Hmm), Active Appearance Model (AAM)-2D Morphable and 3D Morphable Model [18]. Finally, the is also neural network approaches which try to use the power of recent artificial neural networks to find and intuitive representation of faces. some famous ones are Facenet model [19], VGG Google model [20], OpenFace model [21] and MobileNet model [22]. More details on face recognition can be found in [18], [20].

3. PRESENTATION OF FaReS

3.1 Introduction

In this section, presentation ofFaReSapproach will be made. As already explained its merits is the inclusion of moving objects detection and image enhancement in the pipeline of face recognition.More details on the whole process will be brought on the following. The approach is basically made up of eight (08) steps as opposed to four (04) steps in the normal approach. In the normal approach, the first step is input image pre-processing which can include filtering, resizing, etc., followed by a face detection approach, features extraction step and it finishes with face recognition or person identification. This is illustrated on Figure 1. The detection step is there to detect the object to be analyzed or identified. This may include a person, a car, an animal, an area, etc. In the case of this study, emphasis is made on face detection because the application is used in video surveillance.

So, the second step will consist of detecting faces using one of the approaches found in the literature [15], depending on the specificity of the application proposed and the objectives to attain.

The final step consists of the recognition intending to identify the objects detected, that is, to give the nature, the name, the specie, etc. of the detected object. This step is also known as the classification step. In the case of video surveillance for persons, the recognition step will consist of the identification of persons. In other words, it will consist of giving the name of the person detected.

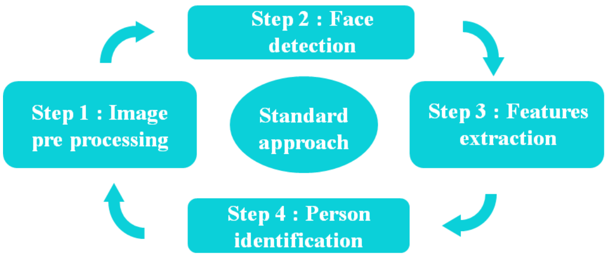

Fig. 1. Traditional steps in face detection and recognition

3.2 Motivation on the choice of the present model

To process multimedia information, definitely more resources than the normal. This increases the complexity of the related application and this complexity gets bigger as the application gets more sophisticated. Consequently, it will be wise to propose solutions that will reduce the mentioned resources consumption. In the specific case of video surveillance, It has been thought wise to first of all detect the presence ofobjects capturing the domain of interest (moving object in this case) before launching post processing on a crop input frame obtained after pixel segmentation.

In addition, it is usually assumed that input images are of good quality while undergoing processes which is not always the case in practice. Consequently, it is advised to insert a step that will be responsible of enhancing the image when it does not satisfy the awaited conditions. This step refers to the resolution in the case study; so, if the input image (face) has a low resolution it must be enhanced so that the new resolution will be greater than the previous one. This gives rise to the architecture presented in the following section.

3.3 Architecture of the proposed approach

Having presented the different problems faced by existing models, let us present the principle of FaRES approach. This approach is no more respecting the conventional paradigm of person recognition which is image pre-processing, face detection and person identification, many steps have been added. These steps include motion detection, foreground extraction and image enhancement. The architecture of the approach is presented on Figure 2.

Note that modification has been done compare to the architecture presented in [3]. Initially, image enhancement was performed before face detection, but in the practice,it has been noticed that in some cases very low resolution (30 × 30 to 08 × 08 pixels) faces are extracted implying a drop in recognition performance. So, to avoid increasing the complexity of the approach with a second process of image enhancement an insertion of the giving process is made after face detection. This enables us to come out with aspects that are being influenced by the approach notably the system memory usage, the image size examined, the detection ratio and the recognition ratio. Let us summarize these steps:

- Step 1 Pre-processing: As in the normal scenario it consists of all the pre-processing mechanisms which could include image acquisition, image denoising, etc.
- Step 2 Motion detection: which consist in the detection of moving objects since the study is made

in the context of video surveillance, the idea is to focus just on the moving object in forget the rest of the image. The approach used here is a background subtraction approach which have been proposed in [28].

- Step 3 Decision making: which consist of stopping the whole process if there is no object to examine and go back to step 1 this helps to save computational resources. This is followed by segmentation and extraction incase moving objects are found; extraction of the image zone of interest is made.
- Step 4 Face detection: Now that the image (frame) is gotten, the algorithm continues by proceeding to face detection to find the human face in the image. The approach used here is MTCNN approach which consists of a cascade of neural network to efficiently detect a face.
- Step 5 Second decision making: Here, a checkup is made to see if the detected face has a sufficient resolution, if the answer is"yes", it is followed by features extraction, otherwise an enhancement is applied.
- Step 6 Image enhancement: Here the idea is to increase the resolution of the detected face if it is less than a given threshold because an image with a small resolution will significantly affect the result. This threshold is chosen empirically depending on the type of application in which it is used and depending on the resolution of the video sensor (Camera). In this case, after many tests, a reasonable resolution has been obtained which is 25×25 pixels. Also, in this study, the choice of one of the famous approaches known as image super-resolution [23], [24] is made. This approach uses a convolution neural network to increase image dimensions. It has been trained to produce a model that multiplies by 4, the size of the input image. Since the complexity of the super resolution in terms of time and memory are negligible compared to the other processes, it does not affect that the much the overall complexities. In the present implementation for example, the time taken is less than a second and the memory used is equivalent to a few bytes on a personal laptop with minimal characteristics. This implies a linear complexity of $\theta(\epsilon)$ in both cases where ϵ tend to 1. In addition, the factor 4 has been determined empirically; meaning significant results have been obtained when the factor was 4. Consequently, multiplication factor can be any factor equal to or greater than 4 depending on the accuracy targeted.
- Step 7 Features extraction: After the face has been detected, a neural network to extract the face features is used. It has been proven in [25] that VGG model tend to be more efficient in terms of accuracy of results. But rather than using the model for recognition it is instead used to extract a feature vector of size 2622. More details on VGG model are found in [26].
- Step 8: Finally, recognition of the person detected is performed. This consists in finding the features vector that is closest to the detected and extracted features vector. In this view, a softmax [27] layer have been trained and added to the previous VGG model.

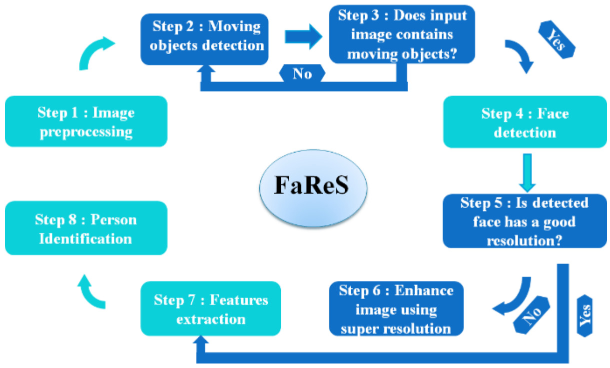

Fig. 2. General architecture of proposed approach

3.4 Brief summary of EFF-ViBE principle

The EFF-ViBE approach as any background subtraction algorithm has three parts namely initialization, segmentation and update which are going to be briefly explained in the followings.

3.4.1 *EFF-ViBE Background model initialization*

To initialize the background model, rather using the initial frame, the first n frames are instead use to compute a mean value and add to the normal N sample values of the normal ViBE algorithm, to have $N + 1$ sample values for each pixel. So, a pixel will now be modeled by

$$M = \{v_0(x), v_1, v_2, \ldots v_N\} \quad (1)$$

Where

$$v_0(x) = \frac{1}{n}\sum_{t=1}^{N} v^t(x) \quad (2)$$

Also, an adaptive radius denoted*Rad* is used in order to take into account the great variability of the environment. So *Rad* is function of the normal R (which is constant) with the following formula:

$$R_{ad} = \begin{cases} R \times (1+\varepsilon_c), & if\ R \le d_{mean} \times \delta \\ R \times (1-\varepsilon_d), & else \end{cases} \quad (3)$$

where parameter $\delta = 6$ rather than 5 (this has been determined empirically) and

$$d_{mean} = \sum_{i=1}^{N} D^i(x) \quad (4)$$

Where $D^i(x)$ is the distance between pixel x intensity value $v(x)$ and sample valuev_i. More details can be found in [28]

3.4.2 *EFF-ViBE Background Segmentation*

If Bdenotes bythe segmented image, $B(x)$ will be the value of the pixel x in matrix B which is equal to 0 if x is a background pixel and 255 if x is a foreground pixel. Therefore, to classify a pixel as background pixel or foreground pixel the following equations are used:

$$R_{ad} = \begin{cases} 0, & if\ R \ge U_{min}\ \ and\ \ v^t(x) \in S_T(v_0) \\ R \times (1-\varepsilon_d), & else \end{cases} \quad (5)$$

$$U = \left|S_{R_{ad}}(v^t(x)) \cap \{v_1, v_2, \ldots v_N\}\right| (6)$$

Where $v^t(x)$ is the value of pixel *x* at time *t*

$$T = \begin{cases} R_{ad}, & if\ \sigma \leq \beta \times R_{ad} \\ \sigma/\beta, & else \end{cases} \quad (7)$$

$$\sigma = \sqrt{\sum_{t=1}^{n}(v^t(x) - v_0(x))^2} \quad (8)$$

Values of all parameters are to find in the work [28].

3.4.3 *EFF-ViBEBackground model update*

To update the model apart from using the mechanisms of the native ViBE algorithm, a pixel counting mechanism is added. This mechanism classifies any pixel that stays in foreground within *K* consecutive frames as background pixel. A matrix with the same size of the frame is maintained and its values represent the number of times a pixel appears as foreground pixel during *K* consecutive frames. If *B*(*x*) denotes the value of pixel *x* and *count*(*x*) its corresponding counter knowing that *B*(*x*) is equal to 0 if *x* is a background pixel and 255 if it is a foreground pixel, *B*(*x*) and *count*(*x*) are updated using the following equations.

$$B(x) = \begin{cases} 0, & if\ count(x) > counter_{max} \\ B(x), & otherwise \end{cases} \quad (9)$$

$$count(x) = \begin{cases} count(x) + 1, & if\ B(x) = 255\ and\ count(x) \leq\ counter_{max} \\ 0, & if\ count(x) > counter_{max} \end{cases} \quad (10)$$

3.4.4 *Implementation*

Experiments have been carried out and the proposed approach has been simulated on the well Known CDnet dataset and compared with some popular background subtraction algorithms and a summary is presented in Table I and more details can be found in [1]. It has been noticed that the proposed EFF-ViBE approach outperforms the native ViBE algorithm in all scenarios. The same observations can be made with other algorithms except SUBSENSE. This can be justified with the fact that even if it outperforms EFF-ViBE in some cases it is relatively slow, and therefore not suitable for real time applications.

3.5 Segmentation and Extraction

Once moving objects have been detected, a circumscription of the zone of interest I made which will consist in the future to the new input image. For that, the set of foreground pixels detected by the segmentation issued from EFF-ViBE is used. At this stage,there is a binary matrix with the size of the input frame where foreground pixels are inwhite (255) and background pixels are in black (0). Then, the area occupied by those foreground pixels is computed followed by edge detection of this area. Once this zone has been restricted, any pixel within is considered as pixel of interest. Therefore, an extraction of this zone of interest is made using cropping mechanism which will have in worse the same size as the input frame otherwise it will have a reduced resolution.

Table I:recapitulating table of f-measure; KDE: kernel density estimation, MoG: mixture o Gaussian, Subsense: self-balanced local sensitivity vibe: visual background extraction, eff-vibe: efficient vibe

Category	Codebook [	KDE	MOG	SUBSENSE [	ViBE	EFF-ViBE
Shadow	0.53	0.53	0.66	<u>0.69</u>	0.63	0.67
Blizzard	0.54	0.52	0.57	<u>0.59</u>	0.55	<u>0.59</u>
Baseline	0.48	0.51	0.64	<u>0.68</u>	0.65	0.67
Camera Jitter	0.54	0.51	0.54	<u>0.58</u>	0.57	0.57
Dynamic Background	0.45	0.51	0.56	<u>0.61</u>	0.58	0.59
Intermittent Object	0.54	0.51	0.55	0.54	0.56	<u>0.59</u>
Night Videos	0.52	0.49	0.51	0.52	0.52	<u>0.53</u>
Low Frame Rate	0.54	0.50	0.53	0.51	0.54	<u>0.54</u>

4. RESULTS AND DISCUSSION

To study the impact of background subtraction and image super-resolution in this approach the following actions have been made:

4.1 Study on the impact of Background Subtraction

First, a video of one minute and few seconds has been considered where the application of EFF-ViBE segmentation approach combine to a developed cropping approach has been made.

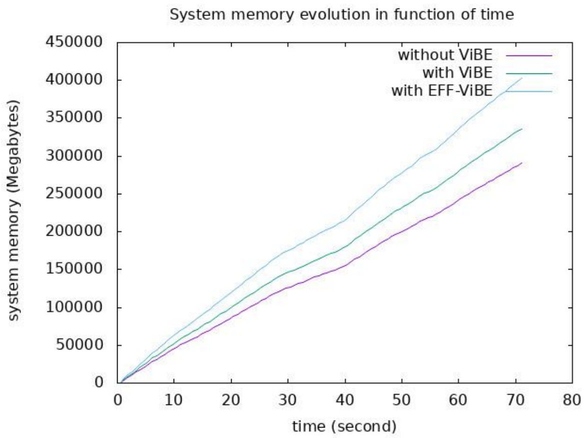

Fig. 3. System memory evolution in function of time

This cropping approach crops the zone of interest in the frame which is the part containing the moving object. It has then been noticed that in terms of system memory usage, the use of background subtraction approach does not significantly affect the memory used for the whole process because the approach is light. This is illustrated on Figure 3 where one can see the evolution of memory consumption with and without background subtraction and added to that the memory usage when using EFF-ViBE. Additionally, it can be noticed that it is better to use EFF-ViBE rather than the normal ViBE segmentation approach.

On the other hand, a record of image sizes being processed during the operation has been made. This has enabled to notice that when applying background subtraction approaches notably ViBE and EFF-VIBE the size of the image treated is lesser than the one used without background subtraction approach.This is understandable in the sense that just part of the image is being processed. An illustration is made on Figure 4.

Fig. 4. Moving Object Detection and extraction, column A: input frame, column: B moving object extraction using ViBE, column C: moving object using EFF-ViBE

On this figure, one can see that image cropped when using EFF-ViBE is better than the one cropped when using ViBE. Also Figure 5 presents the evolution of image size and it can been seen the significant effect of the use of background subtraction approaches. This is logical because without background subtraction the whole frame (image) is considered

4.2 Study on the impact of image super resolution

A model that multiplies by 4 the size of the input face image has been trained; the evolution of the loss function is function of the number of epochs which is 100. With this trained model, image resolution enhancement has been made possible.

For demonstration purposes,400 images of 20 individuals (20 per person)have been used for training, 60 for validation (3 per individual). For testing, 200 images (10 per individuals) have been downscaled with the following resolutions: 25×25, 35× 35, 45×45, 55×55. The idea was to show how image enhancement improves face detection and recognition.

As results, the summary of the number of faces detected as well as the corresponding percentages with theaforementioned resolutions are reported in Table II. On the other hand, the average percentages of good recognition of individuals taken at the testing phase are reported in Table III.

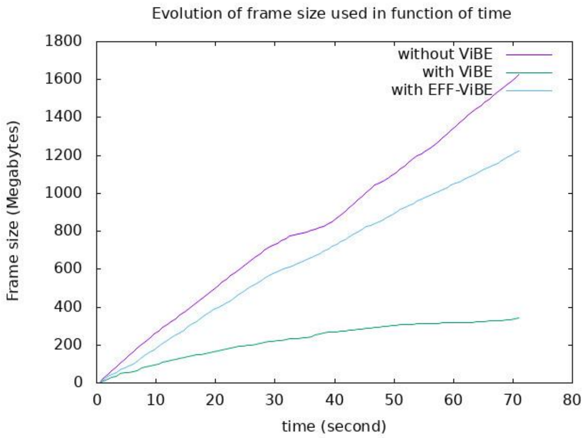

Fig. 5. Evolution of image size used in function of time

TABLE II
NUMBER OF FACES DETECTED IN FUNCTION OF THE INPUT IMAGE RESOLUTION

Resolution	25×25	35×35	45×45	55×55
Normal Face detected (on 200)	175	189	189	191
Enhanced Faces detected (on 200)	191	195	195	197

It can be noticed that low image resolution leads to poor results in the face detection phase as well as at the recognition phase. As image resolution increases, a direct impact on the performances is also observed.

TABLE III
AVERAGE PERCENTAGE OF GOOD RECOGNITION IN FUNCTION OF THE INPUT IMAGE RESOLUTION

Resolution	25×25	35×35	45×45	55×55
Normal Face detected (on 200)	5	36	63.5	79
Enhanced Faces detected (on 200)	32.5	58	76	88

In addition, observing values presented in the two tables, super-resolution mechanism has a good influence on the results in the sense that it significantly increases the number of faces detected as well as the recognition percentage

5. CONCLUSION

This paper has presented an approach called FaReS: A face recognition system based on motion detection and image super resolution. This approach focuses not only on traditional steps in face detection and recognition but includes a moving object detection which reduces the amount of data to be processed as well as an image enhancement increasing recognition accuracy. In this line, related work on moving objects detection, face detection and facerecognition has been presented, later on, detailson FaReS approach explained followed by some implementations proving the efficiency of the approach on resources consumption (especially memory) as well as enhancement of recognition accuracy. However, the proposed system still has some drawbacks because it does not take into consideration partial faces and is focused only humans. Therefore,further studies will first of all include partial faces handling and later on investigations will be made to take into consideration multi-object.

6. REFERENCES

[1] O. Barnich, M. V. Droogenbroeck, Vibe: A powerful random technique to estimate the background in video sequences, International Conference on Acoustics, Speech, and Signal Processing (ICASSP)(2009) 19–24

[2] Thierry Bouwmans, Recent Advanced Statistical Background Modeling for Foreground Detection: A Systematic Survey, Computer Science Review, vol. 4 pp 147-176 (2011) [13] Chris Stauffer, Eric Grimson, Adaptive background mixture models for real-time tracking, Computer Vision and Pattern Recognition, pp 246-252 (1999)

[3] Lionel L. Sop Deffo, Elie Fute T., Emmanuel Tonye, ”LIFADER: Light Face Detection and Recognition approach for people tracking”, Proceeding in IEEE 2nd International Conference on Electronics, Control Optimization and Computer Science (ICECOCS), Kenitra, Morocco on December 2-3, 2020.

[4] Bo Han, Xinggang Lin, Update the GMMs via adaptive Kalman fltering, International Society for Optical Engineering, pp 1506-1515 (2005)

[5] Yang Hong, Yihua Tan, Jinwen Tian, Jian Liu, Accurate dynamic scene model for moving object detection, International Conference on Image Processing (ICIP), pp 157-160 (2007)

[6] Wei Zhang, Xiangzhong Fang, Xiaokang Yang, Jonathan Wu , Spatio-temporal Gaussian mixture model to detect moving objects in dynamic scenes, Journal of Electronic Imaging, (2007)

[7] Jong Geun Park, Chulhee Lee, Bayesian rule-based complex background modeling and foreground detection Optical Engineering, Optical Engineering, (2010)

[8] Thierry Bouwmans, Recent Advanced Statistical Background Modeling for Foreground Detection: A Systematic Survey, Computer Science Review, vol. 11 pp 31-66 May (2014)

[9] BowmaZhihao Wang, Jian Chen, Steven C.H. Hoi, ”Deep Learning for Image Super-resolution:A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence PP(99):1-1 March 2020.

[10] Cui-huan DU, Hong ZHU, Li-ming LUO, Jie LIU, Xiang-yang HUANG, Face detection in video based on AdaBoost algorithm and skin model, The Journal of China Universities of Posts andTelecommunications,Vol. 20 pp 6-9 (2013)

[11] Yi-Qing Wang, An Analysis of Viola-Jones Face Detection Algorithm, Published in Image Processing On Line, pp:128-148, August 31, (2013)

[12] YuseokBan, Sang-KiKim, SooyeonKim, Kar-AnnToh, Face detection based on skin color likelihood, Pattern Recognition, vol. 47, pp: 1573-1585 (2014)

[13] Mohammed Javed, Bhaskar Gupta, Performance Comparison of Various Face Detection Techniques, International Journal of Scientifc Research Engineering Technology (IJSRET), pp 019-0027, Vol. 2 April (2013)

[14] Henry A. Rowley, ShumeetBaluja, Takeo Kanade, Rowley Neural Network-Based Face Detection, IEEE PAMI, (1998)

[15] Kumar, A., Kaur, A. and Kumar, M. Face detection techniques: a review. ArtifIntell Rev 52, 927–948 (2019).

[16] Peter N.Belhumeur, JoaoP.Hespanha, David Kreigman, Eigenfaces vs. Fisherfaces Recognition using class specifc Linear Projection, IEEE Trans. PAMI, (1997)

[17] M. Turk and A. Pentland, Face recognition using eigenfaces, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (1991)

[18] T. Kanade, Computer Recognition of Human Faces, Basel and Stuttgart Birkhauser, (1977)

[19] Florian Schroff, Dmitry Kalenichenko, James Philbin, FaceNet: A Unifed Embedding for Face Recognition and Clustering, arXiv (2015)

[20] O. M. Parkhi, A. Vedaldi, A. Zisserman, Deep Face Recognition , British Machine Vision Conference, (2015)

[21] B. Amos, B. Ludwiczuk, M. Satyanarayanan, Openface: A general-purpose face recognition library with mobile applications, CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., (2016).

[22] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, Mobilenets: Effcient convolutional neural networks for mobile vision, applications. arXiv preprint (2017)

[23] Zhihao Wang, Jian Chen, Steven C.H. Hoi, ”Deep Learning for Image Super-resolution:A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence PP(99):1-1 March 2020

[24] Chao Dong, Chen Change Loy, KaimingHeXiaoou Tang, ”Image super-resolution using deep convolutional networks”, IEEE Transactions on Pattern Analysis and Machine Intelligence 38(2) 2014

[25] Lionel L. Sop Deffo, Elie Fute T, Emmanuel Tonye, ”CNNSFR: A Convolutional Neural Network System for Face Detection and Recognition” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 12, 2018

[26] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In 3rd International Conference on Learning Representations, 2015

[27] Shengwei Zhou, Caikou Chen, Guojiang Han, Xielian Hou, ”Double Additive Margin Softmax Loss for Face Recognition”, Applied Sciences 10(1):60, December 2019, DOI: 10.3390/app10010060

[28] Elie Fute T., Lionel L. Sop Deffo, Emmanuel Tonye, ” EFF-ViBE: An Effcient and Improved Background Subtraction Approach based on ViBE”, International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.11, No.2, pp. 1-14, 2019Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.