Face Verification in Uncontrolled Environments for Access Control

Daniel Lopes*, Ricardo Ribeiro*, António J. R. Neves*
*Institute of Electronics and Informatics Engineering of Aveiro
University of Aveiro
3810-193 Aveiro, Portugal
Emails: {lopesdaniel, rfribeiro, an}@ua.pt

***Abstract*—In the past few years, face recognition has received great attention from both research and commercial communities. Areas such as access control using face verification is dominated by solutions developed by both the government and the industry. In this paper, a face verification system is presented using open source algorithms for access control of large scale events under unconstrained environments. From the type of camera calibration to the algorithms used for face detection and recognition, every stage has a proposed solution. Tests using the proposed system in the entrance of a building were made in order to test and compare each solution proposed.**

Keywords–Face Recognition; Face Detection; Access Control; Unconstrained Environment; Camera Calibration.

I. INTRODUCTION

As one of the most successful applications of image analysis and understanding, face recognition has recently received significant attention and many new techniques have been developed, especially during the past few years [1].

Most face recognition techniques have been developed to be implemented in biometric-based systems and appears to offer several advantages over other biometric methods. An important advantage appointed by [2] regarding these type of systems is the lack of interaction of the user. In a fingerprint system, for example, the user needs to place his finger in a designated area while in a face recognition system the face images can be acquired passively.

Areas related with security, surveillance, access control and multimedia management are some of the fields with an increase demand of face recognition systems. However, there are some levels of complexity regarding these systems as there are some stages that are needed to execute in order to achieve a system with a good performance. These stages are presented in Figure 1.

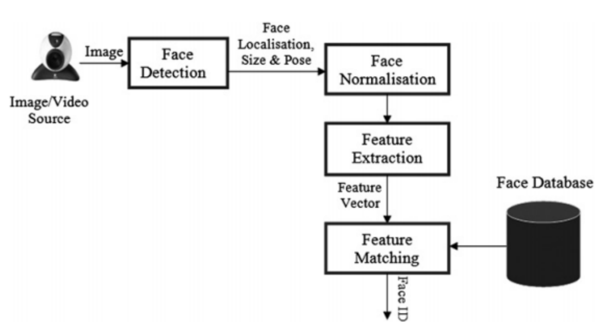


Figure 1. Configuration of a generic face recognition system. [3]

Within each stage, there are specific operations that can be added in order to achieve better performance results. Right on the start, the image acquisition is a crucial step where there is room for improvement. Later, the face detection and recognition can be performed by some specific algorithms, which are presented and studied. Finally, two of the face normalization (also known as preprocessing) algorithms, which are mentioned on state of the art articles are also analyzed for this specific project.

State of the art face recognition systems are dominated by industry and government using large scale datasets. There is a large accuracy gap between todays publicly available face recognition systems and the state of the art private face recognition systems [4]. However, this gap is closing up as it starts to appear better open source algorithms and datasets with more and better images.

Despite the success and high verification or recognition rates, there are still some challenges such as age, illumination and pose variations. Most of these systems work well under constrained conditions (i.e., scenarios in which at least a few of the factors contributing to the variability between face images are controlled), however the performance degrades rapidly when they are exposed to conditions where none of these factors are regulated [2].

In this paper, towards exploring this field, an access control solution for unconstrained environments is proposed using face recognition with open source algorithms. An introductory section is presented that provides a brief introduction to the face recognition system. In Section II, the proposed solution is described. Later in this section, the major problems for a face recognition system for unconstrained environments is explained. These problems are some of the challenges that will achieved and solved in this paper. Section III presents the hardware used in the system. The several implemented algorithms are described in Section IV. In Section V, there are provided experimental results showing the effectiveness of the proposed algorithms and the comparison between them. Finally, a summary of the work done, comparison of the different experiments, concluding remarks and the future work are featured in Section VI.

II. PROPOSED SOLUTION

The project consists in the creation of a face verification (1:1 match comparison) system using open source face recognition and detection algorithms. The main goal is the implementation of this system in large-scale events with access control, such as sport infrastructures. An example of people accessing a sports infrastructure is presented in Figure 2.

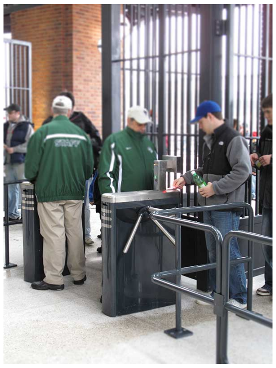


Figure 2. People accessing a sports infrastructure through turnstiles with manual security check. [5]

To access this type of events, it is usually through the acquisition of a ticket/ID card. In order to improve the access control, the ticket access/acquisition are complemented with a face verification system. In the project presented, at the time of the acquisition of the ticket, face images of the person are acquired and then sent to the database. When the person attempts to enter the infrastructure through a turnstile, another face images are acquired and compared with the images that were taken previously.

The environment of these places are usually outdoors, therefore the lighting conditions cannot be fully controlled. In the light of this, a camera parameters calibration is proposed for industrial cameras, which do not have a proper calibration for these type of environments. It was also added an artificial light, which helps to compensate the lack or the excess of light in the scene.

The privacy of people using face verification systems is an important factor in their implementation. The proposed system must be optional, only those who wish to participate, or be mandatory by law in the access control areas.

In order to build a face verification system with these characteristics, an important factor is taken into account: the unconstrained environment where the system is going to be implemented. In a recognition point of view, there will appear some problems related with these kind of environments that are mentioned below.

- **Head pose:** At the time of the image acquisition, the viewing direction of the subject may not be towards the camera. These face images may not be the best suitable for the face recognition system.
- **Face Image Resolution:** As the subject approaches the camera, which he/she is still a few meters from the turnstile, his/her face starts to be detected and tracked. However, if the person is still at some distance from the camera, the face images collected may not have enough resolution for the system.
- **Subject Motion:** It is taken into consideration that the subject is in movement and that may cause some blur in the images acquired.
- **Face Tracking:** It is crucial that there will be distinction between different subjects specially at the time of the ticket acquisition as if not done correctly, the face images of different subjects may end up in the same person database.
- **Non-Controlled Illumination:** This may be the most difficult challenge to overcome as the cameras may be installed in an outdoor environments and, therefore, different lightning conditions according to the time of the day and the meteorological circumstances.

All of these challenges are taken into account when choosing all of the hardware and software for this system.

III. Hardware

In this work, there are used two cameras: the *UI-1220LE-C* (Industrial Camera) and the *Logitech C310* (Webcam). The purpose of the use of these cameras is to compare the performance between them in this specific system as the Webcam does not allow to change its camera parameters such as exposure or gain. On the other hand, the industrial camera, despite not being the most suitable for this scenario, it provides a Software Development Kit (SDK) that enables the fully control of its different parameters. Additionally, as the industrial camera does not have a lens integrated, an 4,5mm lens with manually adjustable aperture is used.

Finally, an 168 LED illumination with adjustable intensity is used in order to compensate the excess or lack of illumination. It also eliminates any occlusion that may be caused by external lightning. Another major advantage is its use on darker scenarios where the camera will have a substantial exposure time. If the illumination is turned on, the scenario will have more light and the exposure needed will be lower thus, the blur captured by the person motion in the image will be far less than with no illumination.

IV. Software

The software developed obeys to some specific steps, which are exposed on Figure 3. The head pose estimation block is not mentioned as it was explained in a previous work [6].

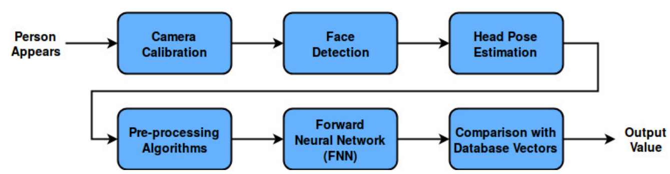


Figure 3. Main steps of the software developed.

A. Proposed Calibration Method

In this section, a different type of calibration is proposed in order to acquire the best digital image for the face verification system.

When using the automatic calibration of the parameters provided by the camera, which is done for the whole image, the Region of Interest (ROI) can be affected by the light intensity that there is in the background.

In order to get a ROI (in this case the face) with the best quality, a calibration focused on this ROI is created. The algorithm proposed is a mixture between the calibration of exposure and gain.

Since the system will be installed in an uncontrolled environment, an initial calibration is done using the auto parameters calibration provided by the camera in order to adapt to the light and environment conditions and to detect the first

 ISBN: 978-1-61208-638-5

face for the use of the proposed calibration. At this point, a timer is set to wait a few seconds, so that the parameters of the camera have time to be internally changed and established. Exposure time, gain and white balance are the parameters changed automatically by the camera software.

Using the auto parameters calibration provided by the camera, the state of art algorithms used for face detection work well for this type of environment. However, this is not true for face recognition algorithms.

When a face is found, the auto parameter calibration is disabled and it continues to the next steps of the calibration.

1) Mean Sample Value: Introduced by [7] in order to create an autonomous setup of the most important parameters of digital cameras for robotic applications, the Mean Sample Value (MSV) is used to set the exposure and the gain of the camera. In this stage, the MSV is calculated through the gray level histogram of the face region with the equation described next.

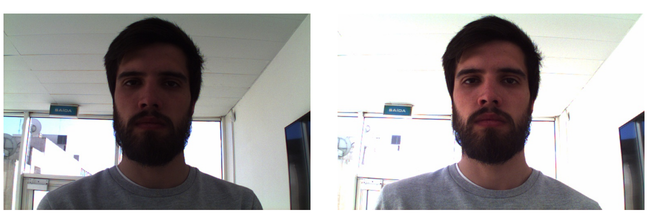
$$MSV = \frac{\sum_{j=0}^{4}(j+1)x_j}{\sum_{j=0}^{4} x_j} \tag{1}$$

where x_j is the sum of the gray values in region j of the histogram (in the proposed approach the histogram is divided into five regions). A range of values is set for the MSV. If the calculated MSV is within that range, the camera parameters (gain and exposure) have acquired values.

This method has the main advantage that, if the same person appears on different parts of the day, the face images acquired will have very similar intensity values as the gain is calculated in order to have the same intensity values between a certain range.

2) White Pixel: This method addresses the situations when the face of a subject is partially directly exposed to sunlight, which causes that part of the face too bright. In order to solve this, if a region where the intensity pixels have the maximum intensity is found, the camera parameters values are decreased in order to reduce the brightness of that region of the face.

Figures 4a and 4b show the comparison between parameters calibration provided by the camera and the proposed calibration, respectively.



(a) (b)

Figure 4. Comparison between the automatic calibration and the proposed calibration.

B. Face Detection and Recognition Algorithms

As for algorithms, there were studied and implemented the following ones into the system. These algorithms are state of the art where the use of neural networks is prevalent, which are trying to close the gap between the performance of commercial and open source of face recognition solutions.

- **Face Detection:**
  - **Histogram of Oriented Gradients (HoG):** *Dlib*'s [8] implementation based on the algorithm presented in [9] that it is used for the face detection stage. Specially useful as it provides 68 face landmarks that are further used at the recognition step for pose estimation.
  - **Multi-task Cascaded Convolutional Networks [10]:** Deep cascaded multi-task framework which exploits the inherent correlation between detection and alignment to boost up their performance. It provides 5 major face landmarks instead of the 64 of *Dlib*. It is, however, more immune to light variations and occlusion.
- **Face Recognition**
  - **Deep Metric Learning (DML):** Implementation also provided by *Dlib* library where the network implemented was inspired in [11] that does the face verification. The model trained achieves an 99.38% in the benchmark Label Faces in the Wild (LFW) [12]. The input data of the network model for training were two datasets: the FaceScrub dataset [13] and the VGG dataset [14] with about 3 million faces in total.
  - **OpenFace [4]:** Face recognition with deep neural networks, which achieves an accuracy of about 92% on the LFW [12] benchmark. The training of the neural network was done with the CASIA-WebFace [15] and FaceScrub [13] containing about 500k images.
  - **DeepFace [14]:** Algorithm inspired in [16] and [17]. The CASIA-WebFace is used on training. In LFW benchmark, it achieves 99.2% of accuracy. The implementation used of this algorithm can be found in [18].

Noteworthy to mention that the *OpenCV*[19] library was used in the image processing and transformation.

C. Preprocessing techniques

1) Gamma Correction (GC): Gamma is a very important characteristic in any digital system. In the world of cameras, it defines the relationship between a numerical value of a pixel and its actual luminance. The GC enhances the local dynamic range of the image in dark or shadowed regions while compressing it in bright regions and at highlights [20]. However, this operation is still affected by some level of directional lightning as pointed by [21].

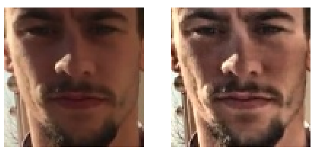
Given a certain gamma (γ), the relation between the gray-level image with gamma correction (I_g) and the original one (I) is given by $I_g = I^{\gamma}$.

Figure 5 presents three images acquired with different gamma values. As it possible to analyze, the image with a higher gamma is more uniform regarding light. The ambition then is that using an appropriate gamma value, the images acquired will not be as susceptible to lightning variations.

Figure 5. Images acquired with gamma of 1, 1.6 and 2 respectively.

2) Contrast Limited Adaptive Histogram Equalization (CLAHE): CLAHE is an adaption of Adaptive Histogram Equalization (AHE) [22] that was first introduced for contrast enhancement for both natural and non-visual images [23].

Figure 6 shows a face image before and after the application of the CLAHE.



(a) (b)

Figure 6. Face image without and with the application of CLAHE

This variation that introduced the limitation of contrast started began to be used in the face recognition field [24], which improved the contrast in face images.

Later, it began to realize its utility in the facial recognition field and a variation entitled of contrast limited adaptive histogram equalization (CLAHE) was started to be used. In this approach, the face image is divided into small blocks, also called tiles, and in each of these blocks the histogram equalization is applied. However, if any of the histograms calculated is above of the predefined contrast limit, the pixels are clipped and distributed uniformly to other bins before applying histogram equalization.

V. EXPERIMENTS

In order to test both software and hardware for the proposed system, an access control system was simulated with face verification at the entrance of the research unit, Institute of Electronics and Informatics Engineering of Aveiro (IEETA), where this work was developed. In these tests, the participation was optional, where the data retrieved from the people remained private.

The system consisted in both industrial and webcam cameras acquiring images with the artificial light, a laptop processing all the software and showing a interactive message to the people who would agree to try to participate in this study. A NFC card reader was also connected to the laptop that would help to register/compare the face images to the NFC number tag provided by the card that the users presented. Figure 7 presents the system set up.

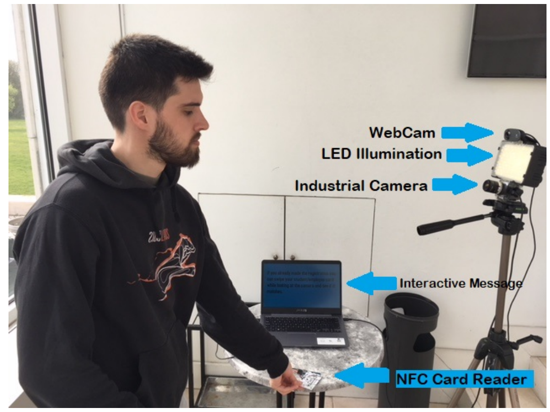


Figure 7. Set up of the system for the experimental results.

The tests were done in three distinct days where the first and the third day were sunny and the second one was cloudy. People who were entering the building were asked if they want to participate this study. If the person agreed, he/she posed himself/herself in front of the camera and the registration was done (if it was the first time that the person presented in front of the camera). As for the next times that the person appeared, the comparison between the face images made on registration and the ones acquired at the time was made. Figure 8 displays some of the face images acquired in the different days.

Figure 8. Example of face images acquired in different days with different meteorological conditions.

About 50 people (a big majority of Caucasians from both sexes) participated and all the participants entered the building at different times of the day, which caused different types of directions of lightning in the face images acquired.

The comparisons between the face images registered in the database and the ones acquired next gave output values, which were used to construct the Receiver Operating Characteristic (ROC) curves. In total, about 2500 comparisons values with both false and true positives were used to construct each curve presented next.

As for the processing times measured, the CPU *Intel Core i7 8550U* was used for the processing of all algorithms.

A. Camera Calibration

The first test analyzes the performance between the webcam with its automatic calibration and the industrial camera with the calibration proposed. Figure 9 presents the ROC curve as well the Area Under Curve (AUC) for this comparison.

The HoG and the Openface algorithms were used with both cameras for detection and recognition respectively.

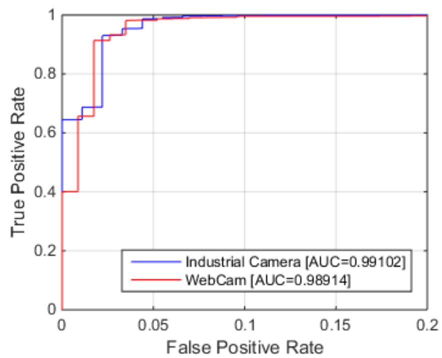


Figure 9. ROC curve comparing the WebCam and the Industrial Camera performance using the same algorithms.

B. Face Detection Performance

In this section, the performance of the HoG and MTCNN face detection algorithms is presented. It was first measured the time that it takes to detect faces in images with dimensions of 752×480 pixels. Posteriorly, the accuracy of each algorithm was tested using a video recorded at the time of the tests. Table I provides the results for both algorithms.

TABLE I. PROCESSING TIMES, TOTAL DETECTIONS AND FALSE POSITIVES FOR EACH FACE DETECTION ALGORITHM.

	HoG	MTCNN
Processing Time (ms)	60	121
Total Detections	592	740
False Positives	1	8

As it can be seen in the table above, both algorithms presented a good performance. The MTCNN algorithm detects more faces, since subjects in profile view are detected.

C. Face Recognition and Preprocessing Algorithms Performance

Results of the performance of the recognition algorithms tested with and without the preprocessing techniques of gamma correction and CLAHE are presented here. As all of the algorithms are based on neural networks it is important to point out that, despite using a specific preprocessing technique, the network was not retrained. The results might improve if the preprocessing technique is applied to the images that are used to train the neural network.

Figures 10, 11 and 12 present the algorithms performance using no preprocessing algorithms and comparing its results with the use of CLAHE and Gamma Correction.

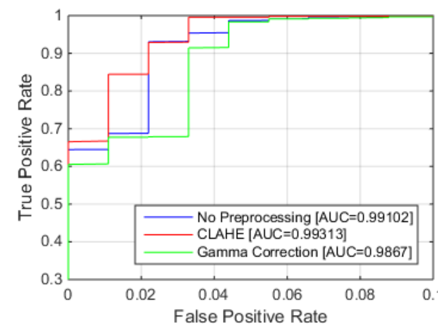


Figure 10. ROC curve presenting the performance of OpenFace using CLAHE, Gamma and no preprocessing technique.

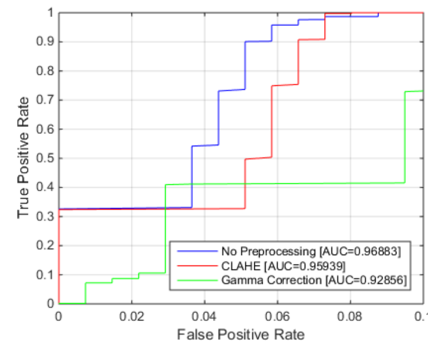


Figure 11. ROC curve presenting the performance of DML using CLAHE, Gamma and no preprocessing technique.

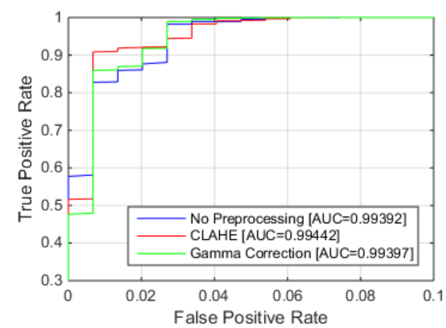


Figure 12. ROC curve presenting the performance of DeepFace using CLAHE, Gamma and no preprocessing technique.

Table II shows the processing time that takes each face image to forward pass the neural network of each algorithm.

TABLE II. PROCESSING TIMES FOR FORWARD PASS IN EACH NETWORK.

	OpenFace	DML	DeepFace
Forward Network Runtime (ms)	236	293	110

The DeepFace is the algorithm with a lower processing time for face recognition.

VI. CONCLUSIONS AND FUTURE WORK

This paper presented a face recognition set up system and studies, which software and hardware is the most appropriate to use under uncontrolled environments. Regarding the camera and its calibration, the industrial camera had a better performance comparing to the webcam as the calibration method presented focus on the best face image that can be acquired. As for software, both detection algorithms presented a good performance. Despite that, MTCNN seems to have the best performance as it detects faces where subject is in the profile view. In relation to the recognition and the preprocessing algorithms, CLAHE algorithm had a positive impact in all of the recognition algorithms as for the gamma correction had a negative impact. It is believed that the results would improve if the preprocessing technique was applied in all of the face images used for the training of the neural network. Unfortunately, the training of these type of neural networks took over a day using powerful GPUs, which are difficult to access. Despite that, the performance overall of the system was satisfactory and, from now on and according to the experiments, the best solution for these type of system is in the use of an industrial camera, MTCNN for face detection, CLAHE for preprocessing and DeepFace for the face verification stage.

The future work goes through the implementation of the system in larger scales where more people would use it. Until then, the train of new neural networks using the preprocessing techniques presented and the study of new alternatives for cameras are on the agenda.

ACKNOWLEDGMENT

This project is partially funded by National Funds through the FCT - Foundation for Science and Technology in the context of the project UID/CEC/00127/2013 and the company ExclusivKey, Lda.

REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM computing surveys (CSUR), vol. 35, no. 4, 2003, pp. 399–458.

[2] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques." Jips, vol. 5, no. 2, 2009, pp. 41–68.

[3] M. Hassaballah and S. Aly, "Face recognition: challenges, achievements and future directions," IET Computer Vision, vol. 9, no. 4, 2015, pp. 614–626.

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

[5] P. TURNSTILES, "Wireless Battery-Powered Portable Turnstiles," URL: http://www.turnstile.com/Product/PWBP/portable_turnstiles_01.jpg [retrieved: April,2018].

[6] D. Lopes, "Pre-processing approaches to improve facial verification in unconstrained environments." RecPad - The 23th edition of the Portuguese Conference on Pattern Recognition, 2017.

[7] A. J. Neves, B. Cunha, A. J. Pinho, and I. Pinheiro, "Autonomous configuration of parameters in robotic digital cameras," in Iberian Conference on Pattern Recognition and Image Analysis. Springer, 2009, pp. 80–87.

[8] "Dlib library," http://dlib.net/ [retrieved: April,2018].

[9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1. IEEE, 2005, pp. 886–893.

[10] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Processing Letters, vol. 23, no. 10, 2016, pp. 1499–1503.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[12] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 2, 2007.

[13] S. W. H. Ng, "A data-driven approach to cleaning large face datasets," 2014 IEEE International Conference on Image Processing, ICIP, 2014, pp. 343–347.

[14] O. M. Parkhi, A. Vedaldi, A. Zisserman et al., "Deep face recognition." in BMVC, vol. 1, no. 3, 2015, p. 6.

[15] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning face representation from scratch," arXiv preprint arXiv:1411.7923, 2014.

[16] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 815–823.

[17] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in European Conference on Computer Vision. Springer, 2016, pp. 499–515.

[18] "Real time deep face recognition," https://github.com/bearsprogrammer/real-time-deep-face-recognition [retrieved: April,2018].

[19] "Opencv (open source computer vision library)," https://opencv.org/ [retrieved: April,2018].

[20] B. Vinothkumar and P. Kumar, "Gamma correction technique based feature extraction for face recognition system," International Journal of Computational Intelligence and Informatics, vol. 3, no. 1, 2013, pp. 20–26.

[21] F. R. Al-Osaimi, M. Bennamoun, and A. Mian, "Illumination normalization for color face images," in International symposium on visual computing. Springer, 2006, pp. 90–101.

[22] K. Zuiderveld, "Contrast limited adaptive histogram equalization," Graphics gems, 1994, pp. 474–485.

[23] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," Computer vision, graphics, and image processing, vol. 39, no. 3, 1987, pp. 355–368.

[24] G. Benitez-Garcia, J. Olivares-Mercado, G. Aguilar-Torres, G. Sanchez-Perez, and H. Perez-Meana, "Face identification based on contrast limited adaptive histogram equalization (clahe)," in Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2011, p. 1.