
Real Time Skin Color Detection Based on Adaptive HSV Thresholding

Mohammed Elamine Moumene*, Khadidja Benkedadra
and Fatima Zohra Berras

Department of Mathematics and Computer Science, Faculty of Exact Sciences and Informatics, University Abdelhamid Ibn Badis Mostaganem, Algeria

E-mail: elamine.moumene@univ-mosta.dz; khadidja.benkedadra@univ-mosta.dz; fatimazohra.berras@univ-mosta.dz

**Corresponding Author*

Received 27 December 2021; Accepted 22 March 2022;
Publication 05 July 2022

Abstract

The detection of human skin color has been studied extensively during the past two decades. It is an essential task for various computer vision applications such as biometric authentication, face/hands tracking and gesture analysis. New machine learning methods are effective for skin color detection. However, they are not suitable for real time applications since they are computationally heavy. A lightweight approach for skin color detection consists of using segmentation rules extracted by an investigation on skin color distribution. The skin appearance varies with diversity of image types, acquisition parameters and scene illumination. There are no general segmentation rules that provide effective skin segmentation for different scene conditions. In this paper we present a real-time skin color detector which adapts itself according to tracked human parts. First, initial thresholds are calculated using two popular skin datasets. Those thresholds can also be calculated quickly using small training sets. The proposed skin color detector

Journal of Mobile Multimedia, Vol. 18_6, 1617–1632.

doi: 10.13052/jmm1550-4646.1867

© 2022 River Publishers

showed comparable skin segmentation to DeepLabV3+ application and an improvement in term of F1 measure when compared to methods that relies on static rules.

Keywords: Human skin detection, skin color, pixel classification, histogram analysis.

1 Introduction to Skin Color Segmentation

Skin color segmentation is the process of separating skin and non-skin regions on images. It is an essential step for various computer vision applications such as face/hands detection and video surveillance. Skin color segmentation is performed on images either by using rules/thresholds or by using machine learning methods. In both cases, skin pixels are classified based on their positions in the chosen color space. The mathematical model defining a color space often uses three variables to represent a specific color. RGB (Red, Green Blue) channels are widely used to store and process digital images. The RGB is an additive color model. It means that different values of Red, Blue and Green light can be used to produce any color. The YCbCr (Luminance, Chrominance) Color Model is a non-linear transformation of the RGB signal. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value. YCbCr values can be obtained from RGB color space according to equations below.

$$Y = 0.299R + 0.587G + 0.114B$$

$$Cb = -0.1687R - 0.3313G + 0.5B + 128$$

$$Cr = 0.5R - 0.4187G - 0.0813B + 128$$

The HSV stands for the Hue, Saturation, and Value. The value represents intensity of a color. The hue and saturation components are inspired from the human eye perception. The transformation equations for RGB to HSV color model conversion are given below.

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

$$S = 1 - 3 \frac{\min(R, G, B)}{R + G + B}$$

$$V = \frac{1}{3}(R + G + B)$$

The three color spaces introduced above are commonly used for skin detection [1, 2]. Albiol et al. [3], demonstrated that every color space contains optimal rules for effective human skin detection. On the other hand, many studies focused on finding the most appropriate color space for an efficient skin detection [1, 2, 4]. HSV and YCbCr correspond to the way that humans describe colors and the separation between chrominance and luminance on those perceptual color spaces makes them suitable for skin color segmentation against illumination variation [5, 6]. In [7], a comparison between HSV based and YCbCr based skin color detection showed that the HSV color space is suitable for simple images, whereas YCbCr produces better results on complex images.

In recent studies, two main approaches are considered for skin colour segmentation: machine learning methods and rule based methods [1, 2]. The first class of methods uses supervised training data to predict pixel's class (skin/non-skin). This kind of methods delivers best results but are computationally expensive. One of the common machine learning methods used for skin color segmentation is SVM [9]. Han et al. [8] proposed a skin detector for hand gesture recognition where a SVM classifier was trained on a skin dataset, then additional region information was used to reduce illumination variations problems. Artificial Neural Networks have also been implemented for the detection of skin color. Kim et al. [10] introduced two convolutional neural networks for skin detection based on different features such as color, texture and shape. DeepLabV3+ is another method based on CNN which achieved the best performances on 10 different datasets [11]. Few studies have approached the skin segmentation problem using convolutional neural network due to the lack of complete datasets [1]. Existing training sets are not representative which leads to the reduction of the segmentation precision [12]. Another serious weakness of using machine learning methods for skin detection is the slow processing time. The second approach for skin color segmentation consists of using rules or thresholds defined according to a color space. Classification rules represent relationships between color components. Several thresholding methods have been proposed in the literature and new algorithms are still being proposed [13, 14]. They are preferred since they are easy to implement, easy to reuse and they are relatively efficient. The rules or

thresholds are made on the basis of skin color distribution observations. For instance, Fernandez et al. [15], the histogram frequencies are transformed into a probability distribution. A pixel is classified as a skin pixel if its likelihood is greater than a predefined threshold. Another fast method for image segmentation is Lookup table [16].

The problem of skin color segmentation turns out to be a difficult problem as the color of the skin may vary in its appearance due to many factors such as illumination variation and complex backgrounds. Color constancy problem influences the most skin detection. Indoor, outdoor, shadows or highlights produce difference in skin color on images. Different camera parameters also produce different color appearances of the same scene under the same lighting conditions. Another factor which can alter human skin detection is a complex background (it may contain objects having skin-like color). Machine learning methods show more efficiency than the rule-based methods. However, they require a long classification time, which is not suitable for real-time applications. Skin segmentation is frequently used as a primary step in other computer vision applications. It is then important for it to be computationally efficient. Rule-based methods are less effective but their main advantage is their simplicity and speed. In this paper we introduce an efficient skin segmentation method based on adaptive thresholds. HSV thresholds for skin segmentation are calculated using two popular and publicly available datasets which are Pratheepan and HGR [1]. After that, motion detection is used to refine thresholds according to human body parts and capturing conditions. A soft segmentation algorithm is considered to satisfy the real time constraint. The remainder of the manuscript is organized as follows: in Section 2 we introduce the proposed skin color segmentation algorithm. The obtained results by the proposed detector and comparisons with some relevant works are described in Section 3. The conclusion is presented in Section 4.

2 The Proposed Adaptive Skin Detector

The flowchart in Figure 1 shows the proposed skin detector algorithm. To define the initial segmentation thresholds, we start by generating a global histogram based on skin color dataset. After that, the segmentation thresholds are estimated based on the histogram's derivatives. Finally, an adaptation phase based on motion detection is used in real-time for skin color segmentation.

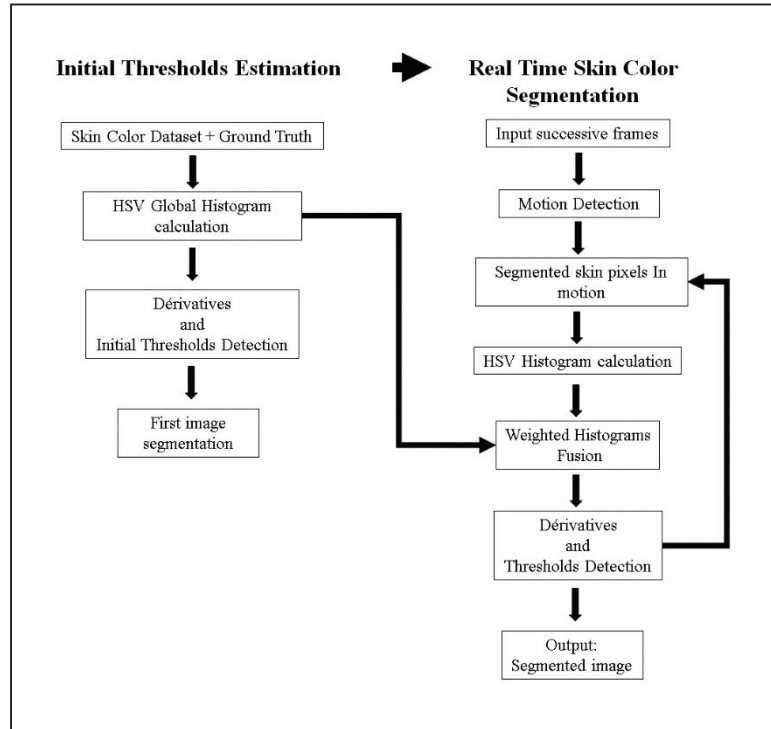


Figure 1 The proposed skin detector based on adaptive HSV thresholds.

One of the best ways to understand data is to plot graphs. The histogram is a useful graphic representation of data, where each color value is characterized by its frequency on the image. Based on the discussion introduced in the previous section, we believe that using histograms to define adaptive thresholding could be one of the fastest methods to implement an efficient skin detector. To define the initial segmentation thresholds we start by generating a global skin color histogram. For that, we use two popular datasets commonly used for skin color segmentation research. The first one is the Pratheepan dataset [1], a small dataset consisting of 78 photos randomly collected from Google. The second one is called HGR. It is a gesture recognition dataset including 1558 images of American and Polish sign language [1]. Both datasets contain ground truth in the form of binary mask characterizing skin pixels. Figure 2 shows samples from both datasets.

The binary masks (ground truth) are used as shown in Figure 3 to create a global list containing all skin pixels contained in a dataset. This list allows



Figure 2 Sample images with their skin colour masks from the HGR (Top) and Pratheepan (Bottom) datasets.



Figure 3 Using dataset masks to extract skin pixels and create global skin Histograms.

the calculation of a global skin histogram on each channel of the three color spaces (RGB, HSV, YCbCr). Afterwards, A Gaussian smoothing and frequencies normalization are applied to the obtained histograms. As shown in Figure 4, distribution of skin color has Gaussian curve form. We can also see that the peaks and widths of the Gaussians on the training histograms are dependent on the datasets. This color distribution difference between datasets

is less important on the HSV color space, which lets us assume that this color space is suitable for universal thresholds estimation.

Finding the skin color thresholds is carried out by Gaussian's parameters estimation which are maximums M and inflection points T . (M_R , T_{Rmin} , T_{Rmax} in the RGB histogram for example). We pose M_R on the maximum found inside the interval $[0, 255]$ of the corresponding Red histograms. Inflection points are those points where the curve of a function changes convexity. The convexity of a function over an interval is related to the sign of the second derivative over that interval. If the second derivative changes sign at a point, then the function changes convexity at that point. Based on these notions, the inflection points T_{Rmin} , T_{Rmax} can be estimated using the first and second derivatives of the smoothed histograms as follows:

$$T_{Rmin} = \min\{i \in [0, 255] \text{ where } H''_R(i) \times H''_R(i - 1) \leq 0 \text{ and } H'_R(i) < 0\}$$

$$T_{Rmax} = \max\{i \in [0, 255] \text{ where } H''_R(i) \times H''_R(i - 1) \leq 0 \text{ and } H'_R(i) > 0\}$$

Where H'_R and H''_R are respectively the first and the second derivatives of the red channel histogram. The same Gaussian parameters are estimated for the green and the blue channels in the same way. Skin segmentation is then performed on input image I using the rule:

$$\text{Segmented Skin Pixels} = \{p_{x,y} \in I \text{ where}$$

$$T_{Rmin} < R_{x,y} < T_{Rmax}$$

$$T_{Gmin} < G_{x,y} < T_{Gmax}$$

$$T_{Bmin} < B_{x,y} < T_{Bmax}\}$$

Where $p_{x,y}$ is a pixel from the input image I having $(R_{x,y}, G_{x,y}, B_{x,y})$ values in the RGB color space. In the next section we present and discuss the calculated thresholds for the three commonly used color spaces: RGB, YCrCb and HSV on both datasets (Pratheepan and HGR).

As stated earlier before, the distribution of the skin color is slightly different between the two datasets (Figure 4). The estimated skin color thresholds are dependent on some factors such as the image capture hardware and the illumination conditions. It is for this reason that it is necessary to add an adaptation phase which takes into account different acquisition conditions. In many skin color based applications such as gesture recognition and surveillance application, the human body parts are often in motion. This assumption leads us to the idea of adapting the thresholds according to skin

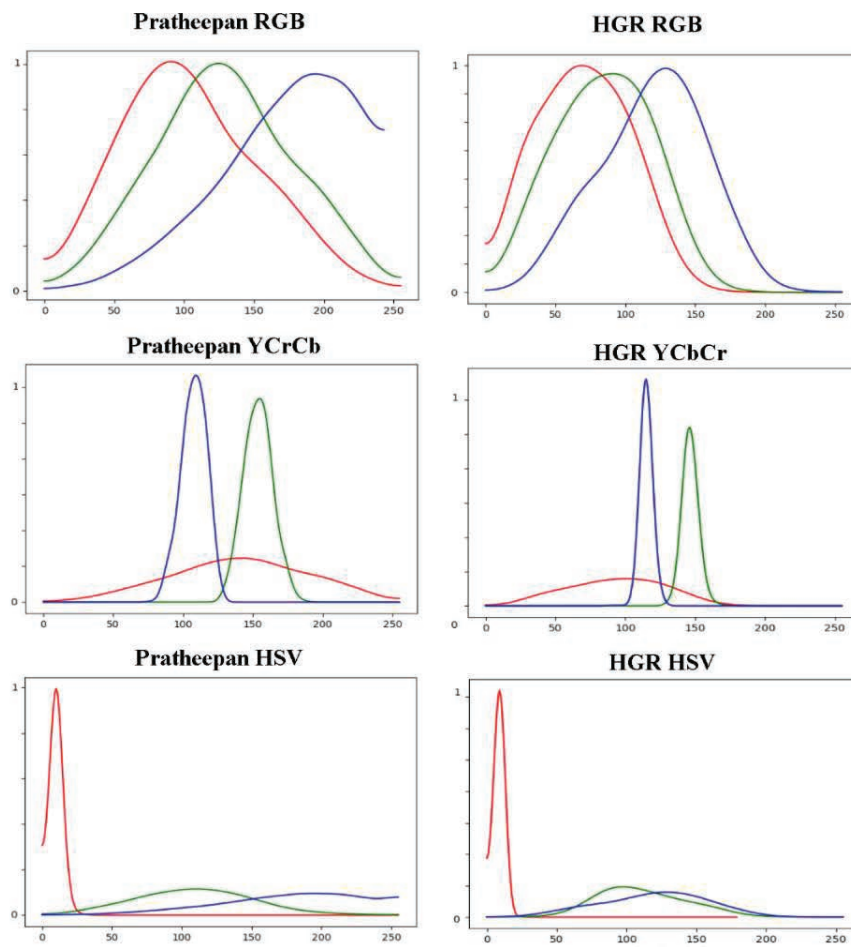


Figure 4 Normalized global skin color Histograms calculated using different color spaces for HGR and Pratheepan datasets.

pixels which are in motion. Once the initial thresholds are estimated from the global skin color histogram, the segmented skin regions are tracked for motion. A normalized histogram is constructed from the segmented skin pixels which are in motion, then fused with the initial global skin histogram using a weighted average. Based on experimentations, we fix the thresholds to 0.3 for the in-motion skin pixels histogram and 0.7 for the Global skin colour histogram. The histogram merging occurs for the successive frames and a simple image subtraction is performed for motion detection.

3 Results and Discussion

Standard measures for performance can be used for skin detection as it is a two-class segmentation problem. We compare results with the ground-truth data using the F1 measure which is a common metric for skin segmentation [1]. The F1 measure is calculated according to the following formula:

$$F1 \text{ measure} = \frac{2TP}{(2TP + FN + FP)}$$

Where skin pixels properly identified as skin are the true positive (TP). False positives (FP) are non-skin pixels falsely segmented as skin. False negatives (FN) are incorrectly identified pixels as non-skin.

We calculate the initial skin color thresholds using the three color spaces RGB, YCrCb and HSV as shown in Section 2. The obtained skin segmentation rules for both experimented datasets are:

- RGB thresholds using Pratheepan dataset:

$$(133 < R < 255, \quad 80 < G < 170, \quad 35 < B < 147)$$

- YCrCb thresholds using Pratheepan dataset:

$$(95 < Y < 187, \quad 138 < Cb < 172, \quad 90 < Cr < 128)$$

- HSV thresholds using Pratheepan dataset:

$$(4 < H < 16, \quad 50 < S < 170, \quad 134 < V < 254)$$

- RGB thresholds using HGR dataset:

$$(95 < R < 163, \quad 15 < G < 171, \quad 50 < B < 88)$$

- YCrCb thresholds using HGR dataset:

$$(60 < Y < 140, \quad 139 < Cb < 153, \quad 60 < Cr < 170)$$

- HSV thresholds using HGR dataset:

$$(2 < H < 16, \quad 45 < S < 149, \quad 60 < V < 198)$$

In Table 1 we report the performance of the skin detectors mentioned above and comparisons with two relevant works in terms of F1 measures and CPU time. These two works are the static rule-based segmentation (SRBS) [13] and the DeepLabV3+ application [12]. The segmentation using the HSV color space showed best results when compared to the segmentation based on the RGB or YCrCb color spaces. The proposed HSV detector produces better results in terms of F1 measures when compared to SRBS method. In terms of processing time, the proposed method outperformed the

Table 1 Objective comparisons between the proposed skin detector and some relevant works

Skin Segmentation Methods		F1 Measure	CPU Time
Static Rules based Segmentation [13]	HGR [1]	0.631	0.10s
Proposed Detector on RGB		0.789	0.30s
Proposed Detector on YCrCb		0.880	0.30s
Proposed Detector on HSV		0.937	0.30s
DeepLabV3+ [12]		0.954	6.27s
Static Rules based Segmentation [13]	Pratheepan [1]	0.503	0.10s
Proposed Detector on RGB		0.541	0.30s
Proposed Detector on YCrCb		0.793	0.30s
Proposed Detector on HSV		0.842	0.30s
DeepLabV3+ [12]		0.875	6.27s

**Figure 5** Visual comparisons using different skin segmentation methods on Pratheepan dataset. (From Left to right): Original image, ground truth, DeepLabV3+ [12], Proposed detector, static rule-based method [13].

DeepLabV3+ approach while achieving comparable F1 measures. The figures 5 and 6 show a visual comparison between images segmented using the different methods. We can see how the static thresholds decrease the number of true positive rates. The proposed adaptive thresholding has a slightly higher number of undetected skin pixels compared with the DeepLabV3+ approach. However, it can detect the actual skin pixels with reasonable rate compared to static rule based method. We can thus conclude that adaptive thresholding is crucial for skin color based applications.

To evaluate the proposed adaptive thresholding we applied it on successive frames captured using a simple webcam. The length of the sequence



Figure 6 Visual comparisons using different skin segmentation methods on HGR dataset. (From Left to right): Original image, ground truth, DeepLabV3+ [12], Proposed detector, static rule-based method [13].



Figure 7 Motion based Adaptive Thresholding. (Left): frame at time t_0 , (Right): Frame at time t_0+5s .

showed in the Figure 7 is 575 frames. We can see that there is a significant enhancement of skin detection precision between the frames t and $t + 5$ seconds. The graphic in Figure 8 demonstrates the improvement of the F1 measure through the successive frames.

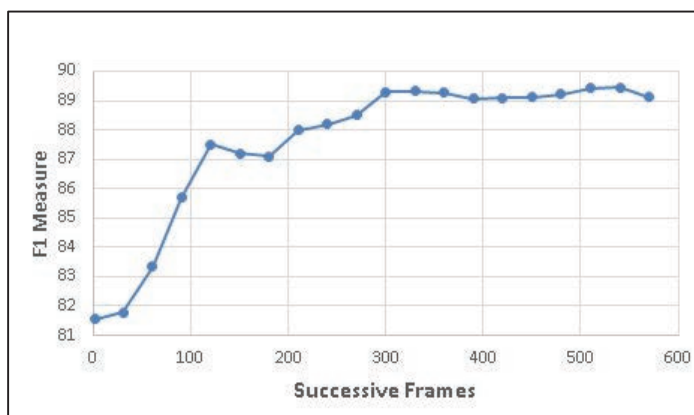


Figure 8 Effect of adaptive thresholding on skin segmentation precision through successive frames.

4 Conclusion and Perspectives

New machine learning approaches such as deep learning are the best solutions for many computer vision applications. In the field of skin segmentation, there is a lack of a universal datasets for the learning process. In addition, machine learning methods are time consuming and therefore not suitable for real time applications. We introduce, in this paper, a lightweight rule-based skin segmentation that can be used for wide range of environments. The proposed system uses HSV thresholds pre-calculated during an initial training. Those thresholds are then recalculated and adapted to the users's skin color. The proposed detector yielded better results in comparison with static rules-based methods and a comparable skin segmentation to the DeepLabV3+ application.

References

- [1] Alessandra Lumini, Loris Nanni, "Fair comparison of skin detection approaches on publicly available datasets", *Expert Systems with Applications*, vol. 160, p. 113677, 2020.
- [2] Sinan Naji, Hamid A. Jalab, Sameem A. Kareem, "A survey on skin detection in colored images", *Artificial Intelligence Review*, vol. 52, pp. 1041–1087, 2019.

- [3] Albiol A, Torres L, Delp E, “Optimum color spaces for skin detection”, Proceedings of the IEEE international conference on image processing, pp. 122–124, 2001.
- [4] Chaves-González JM, Vega-Rodríguez MA, Gómez-Pulido JA, Sánchez-Pérez JM, “Detecting skin in face recognition systems: a colour spaces study”, *Digit Signal Process*, vol. 20, pp. 806–823, 2010.
- [5] Brancati N, De Pietro G, Frucci M, Gallo L, “Human skin detection through correlation rules between the YCb and YCr subspaces based on dynamic color clustering”, *Comput Vis Image Underst*, vol. 155, pp. 33–42, 2017.
- [6] Rahman MA, Purnama IKE, Purnomo MH, “Simple method of human skin detection using HSV and YCbCr color spaces”, IEEE international conference on intelligent autonomous agents, networks and systems, pp. 58–61, 2014.
- [7] Shaik KB, Ganesan P, Kalist V, Sathish B, Jenitha JMM, “Comparative study of skin color detection and segmentation in HSV and YCbCR color space”, *Proc Comput Sci*, vol. 57; pp. 41–48, 2015.
- [8] Han J, Awad G, Sutherland A, “Automatic skin segmentation and tracking in sign language recognition”, *Comput Vis IET*, vol. 3, pp. 24–35, 2009.
- [9] Mehmet F , Utku K, “A novel color-based feature extraction method for svm based skin segmentation”, *eskişehir technical university journal of science and technology*, vol. 21, pp. 45–54, 2020.
- [10] Kim Y, Hwang I, Cho NI, “Convolutional neural networks and training strategies for skin detection”, IEEE international conference on image processing, pp. 3919–3923, 2017.
- [11] Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H, “Encoder-decoder with atrous separable convolution for semantic image segmentation”, *Lecture Notes in Computer Science*, 2018.
- [12] Khan R, Hanbury A, Stöttinger J, Bais A, “Color based skin classification”, *Pattern Recogn Lett*, vol. 33, pp. 157–163, 2012.
- [13] Kolkur, Seema, et al., “Human skin detection using RGB, HSV and YCbCr color models”, *Advances in Intelligent Systems Research*, vol. 137, pp. 324–332, 2017.
- [14] Thwe, Phyu Myo, Nwe Ni Kyaw, and Kyaw Kyaw Naing, “Hand Region Detection using CbCr Color Space and Otsu’s Method”, *International Journal of Trend in Scientific Research and Development*, vol. 3, pp. 1568–1571, 2019.

- [15] Fernandez A, Ortega M, Cancela B, Penedo M , “Human body parts contextual and skin color region information for locating human body parts”; *J Comput Inf Technol*, vol. 1, pp. 1–16, 2012.
- [16] De Siqueira FR, Schwartz WR, Pedrini H, “Adaptive detection of human skin in color images”, *IX workshop de Visão Computacional*, pp. 1–6, 2013.

Biographies



Mohammed Elamine Moumene received a computer engineering degree from Mostaganem University in 2010 and a Ph.D. degree in computer vision from Oran 1 University in 2018. He is currently working as an assistant professor at the Department of Mathematics and Computer Science in Mostaganem university. His research areas include computer vision, machine learning and data mining.



Khadidja Benkedadra received a bachelor degree in Computer Systems from Mostaganem university in 2018, then a Master degree in Information Systems Engineering from the same university in 2020. She is currently working as the IT coordinator of a company and a freelance data analyst. She is specialized in image processing.



Fatima Zohra Berras received a bachelor's degree in Computer Systems from Mostaganem University in 2018, then a Master's degree in Information Systems Engineering from the same University in 2020. She is currently a student in Autonomic Systems at the University of Paris Saclay. She is specialized in image processing and machine learning.

