Redundancy Reduction in Face Detection of Viola-Jones using the Hill Climbing Algorithm

Kartika Candra Kirana
Department of Electrical Engineering
Universitas Negeri Malang
Malang, Indonesia
kartika.candra.ft@um.ac.id

Slamet Wibawanto
Department of Electrical Engineering
Universitas Negeri Malang
Malang, Indonesia
slamet.wibawanto.ft@um.ac.id

Heru Wahyu Herwanto
Department of Electrical Engineering
Universitas Negeri Malang
Malang, Indonesia
heru_wh@um.ac.id

***Abstract*— In this study, we improved Viola-Jones face detection using Hill Climbing algorithm in order to reduce the redundancy rates. Viola-Jones's algorithm has succeeded in determining the best face scale factor, but it produces a non-comparable face window that leads the redundant face detection. Hill Climbing algorithm which is one of the local search family, is proposed to serve the local-maxima which represents a set of faces that has been selected from redundant data. We have evaluated and compared the accurate, accurate and recall tests the performance of the proposed method using LFW dataset. Traditional Viola Jones achieves accuracy up to 77% and the proposed method achieves accuracy up to 85%. The two sets of values concluded that the proposed method reduce the redundancy problem.**

Keywords—face detection, Viola-Jones algorithm, hill climbing, redundant, scale factor.

I. INTRODUCTION

During the Co-19 pandemic that enforced social distancing, the online learning system was encouraged to incorporate facial recognition as a part of the presence and learning emotional recognition [1]–[3]. In addition, facial recognition immersion has been widely used in various systems, such as surveillance systems and biometric systems [4], [5]. The critical element of facial recognition is the face detection. This work is a sub-project that incorporates face detection into online learning systems.

Viola-Jones is the most common algorithm used to detect for faces [5], [6]. This approach refers to the Haar-like feature that is built on the Haar Wavelet [7], [8]. The Haar Wavelet is a single square wave with high and low intervals, then it is computed using an integral principle to generate features in a window. Haar-like features are categorized using the Cascaded Adaboost algorithm to classify face candidates. If the window is the candidate of the face, the scale factor is raised [9]–[11]. The scale factor plays an important role in enhancing the accuracy of the detection. Low-quality image scaling algorithm results in a lack of features that directly affects the performance of the detector. [9]. Various studies have been developed to address the problem of scale factor in the Viola-Jones algorithm. Sharma et al (2015) investigated the image scaling algorithm for Viola and Jones' face detection, which consists of: Bicubic, Bilinear, Extended Linear, Nearest Neighbor, and Piece-wise extended linear [9]. The results show that the Piece-wise extended linear is the best performance relative to all other algorithms [9]. Taloba et al (2018) refines the Viola-Jones scale factor by Genetic Algorithms. The results showed that the scale factor in each database were different in order to achieve perfect area detection [10]. Loresco et al (2018) detects the development of the leaf using the Viola-Jones algorithm by applying a static marker as a standardized area measurement [11]. The accuracy of the leaf scale factor can be detected up to 97.5 per cent [11]. These studies have succeeded in determining the best scale factor in the Viola-Jones algorithm, but the non-comparable windows leads the redundant face detection. The arbitrary results need to be optimized in order to find the better results. Local searches can be introduced to reduce wasteful computing.

In addition, the Hill Climbing algorithm is one of the mathematical techniques used in the local search family. Hill Climbing algorithm has been widely used in the process of image segmentation [12]–[14]. Wei and Yan (2011) segment the brain using Hill Climbing algorithm based on the main curvature of the cortical surface[15]. Ganesan et al (2016) developed the Hill Climbing and Fuzzy C Means algorithm to overcome the ambiguity segmentation of satellite [16], fruit [17], and skin cancer [18]. Roy et al (2016) propose the segmentation of hill climbing based on watersheds to provides a parallel architecture, thus avoiding FIFO queues and saving the memory [19]. Furthermore, the Hill Climbing algorithm has been used in the method of facial verification[20], [21]. These studies show that the Hill Climbing algorithm can be used to separate different segments and to group related segments. The idea of grouping related segments of face candidate and presenting a set of local maxima as updated solution is proposed to reduce the redundancy rate.

In this study, we improved Viola-Jones face detection using Hill Climbing algorithm in order to reduce the redundancy rate. Viola Jones' algorithm was used to achieve an arbitrary face solution. The Hill Climbing algorithm is tasked to serve a local-maxima that represents a face that has been selected from redundant data. Our contribution is embedding the Hill Climbing algorithm in the face detection of Viola Jones in order to reduce the rate of redundancy generated by the Viola Jones algorithm.

In this study, Section I discusses the background of the study. Section II discusses the proposed methodology. Section III discusses the results and discussion. Section IV discusses the conclusions of this study

II. METHODOLOGY

In this study, we improved Viola-Jones face detection using Hill Climbing algorithm in order to reduce the redundancy rate. The pseudocode of the proposed method is shown in Fig.1. The phase of the proposed method consist of: (a) generating face candidates using viola-jones algorithm and (b) generating local optima on face candidates using the hill climbing algorithm.

978-1-7281-8131-8/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Middlesex University. Downloaded on November 03,2020 at 16:33:53 UTC from IEEE Xplore. Restrictions apply.

```
Input: original test image
Output : image with bounding box
For i ← 1 : sum of image do
 Downsample image
 Call function integral image // Eq.(1)
 // Fig.(3)
 For j ← 1 : sum of sub window do
  For k ← 1: sum of cascaded classifier do
   For l ← 1 :  filter of stage k do
     Fiter window
     Accumulate window's total
   end for
   if window's total not reach threshold then
     reject window as face
  else
     accept window as face
     call function comparison  // Eq. (2)
     call function local_optima // Eq. (3)
     show local_optima
   end if
  end if
 end for
end for
```

Fig. 1. Pseudocode of The Proposed Method

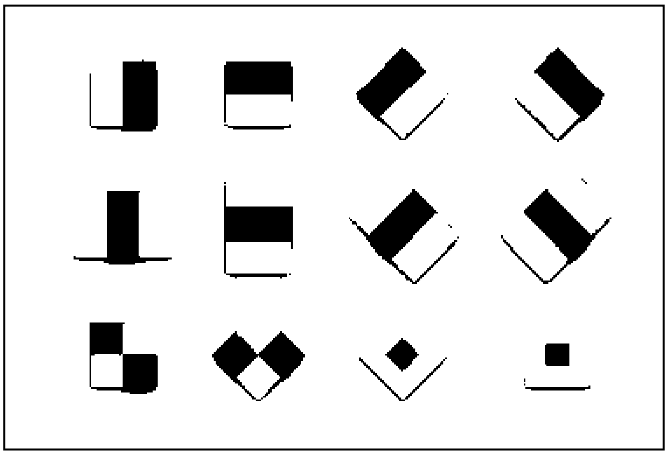


Fig. 2. Haar-Like Feature [11]

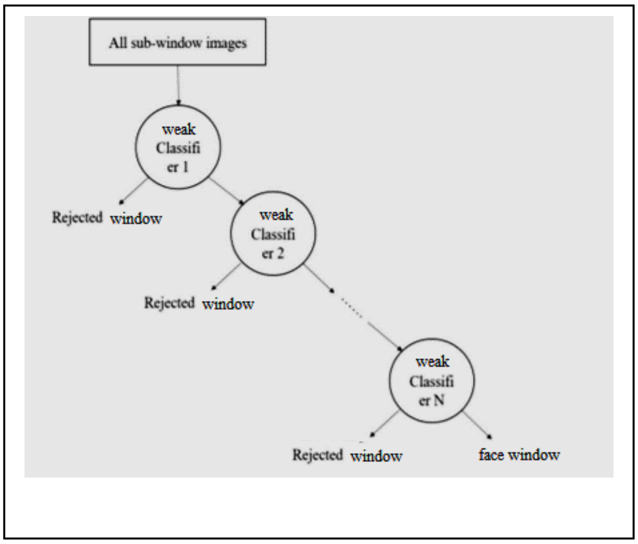


Fig. 3. The Cascaded Adaboost algorithm

A. Generate Face Candidates using Viola-Jones Algorithm

The Viola Jones algorithm applied the Haar Wavelet, which is calculated by applying the pixel strength to the adjacent rectangular window of the detection field. The pixel strength and the threshold are compared in order to determine the appearance of the object in the field. The Haar-like feature is shown in Fig. 1 and is calculated by Equation (1).

$$v = sum(w) - sum(b) \tag{1}$$

The eigen value (v) is the difference value between the pixels of the black (b) and white (w) region.

Haar-Like features are categorized using the Cascaded Adaboost algorithm to classify face candidates. The AdaBoost classification is a strong classifier built by weak classifiers in cascaded decision tree and it is shown in Fig. 3. For this process, a 24x24 pixel window over the entire image from left to right and then from top to bottom. The scale factor is increased once one full scan of the size of this window is done. The output generated at this stage is an arbitrary face candidate. Then the Hill Climbing algorithm is used to find the best face candidates shown.

B. Generate Local Optima On Face Candidates Using The Hill Climbing Algorithm

In a simple hill climb, the first closest node is chosen, while the current node and subsequent nodes are compared, the path of router is compared and the closest to the solution is chosen. The definition of "closest to the solution" is reaching the window similarity threshold and the local maxima is selected using the maximum function. If both nodes fail to reach the window similarity threshold, the current node is considered the comparison node.

The comparison of node is shown in Equation (2).

$$\Delta v = v_x - v_{x-1} \tag{2}$$

The difference of eigen value (Δv) is calculated between the eigen face of current node (v_x) and the eigen face of subsequent nodes (v_{x-1}) . Based on Equation (2), the heuristic value that is processed is the eigenvalue.

Authorized licensed use limited to: Middlesex University. Downloaded on November 03,2020 at 16:33:53 UTC from IEEE Xplore. Restrictions apply.

$$localoptima(x) = \begin{cases} true, & \Delta v \geq \theta \\ false, & \Delta v < \theta \end{cases} \quad (3)$$

III. RESULT AND DISCUSSION

We use a LFW [22] and frontal face cascaded [23] data set to evaluate the performance of the proposed method. LFW dataset is used as the testing data and frontal face cascade is used as the training data. We collected 685 random images with the name tags of 900 people. It means that two or more pictures belong to the same persons in a different background. The sample data is shown in Fig. 4.

In accordance with the purpose of this study to reduce the redundancy of facial detection, we conducted accurate, accurate and recall tests as shown in Equation (4), (5) and (6), respectively.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad , \quad (4)$$

$$precision = \frac{TP}{TP+FP} \times 100\% \quad , \quad (5)$$

$$recall = \frac{TP}{TP+FN} \times 100\% \quad , \quad (6)$$

Based on Equation (4)-(6), True Positive (*TP*) is an multiple face that does not experience redundancy, True Negative *(TN)* is an single face that does not experience redundancy, False Positive *(FP)* is a multiple face image that experiences redundancy, False Negative (*FN*) is the single face image that experiences redundancy.

Fig. 4. The Sample of Data

TABLE I. THE COMPARISON

No	Comparison		
	Test	***Proposed Method***	***Traditional Viola-Jones [2]***
1	True Positive	740	520
2	True Negative	25	0
3	False Negative	100	155
4	False Positive	35	225
5	Accuracy	85%	77%
6	Precision	95%	71%
7	Recall	88%	81%

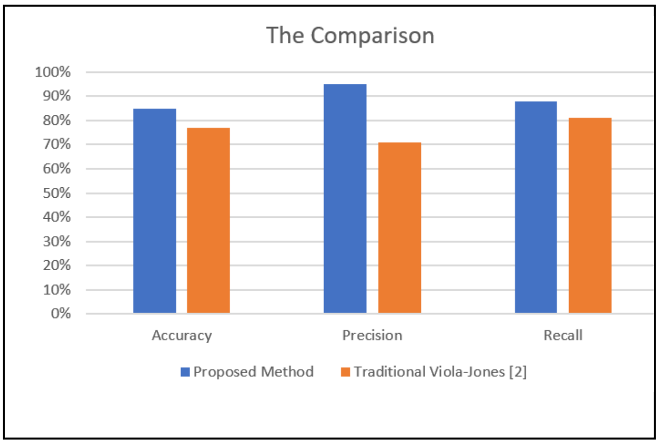


Fig. 5. The Comparison

We compare the series of test between traditional viola jones which is also used in [2] and our proposed method. The comparison result is shown in Table I. The comparison of the Proposed method and the previous method (Traditional Viola-Jones) is also outlined through the graph in Fig. 5.

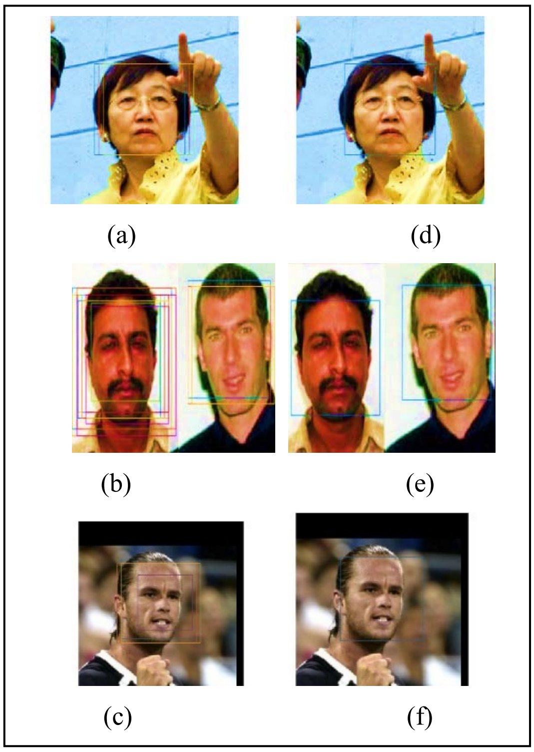
Based on Table I and Fig. 5, there are improvements to the results on precision, recall, and accuracy. Traditional Viola Jones achieves accuracy, precision, and recall respectively at 77%, 71%, and 81%. Meanwhile, the proposed method achieves accuracy, precision, and recall by 85%, 95%, and 88%, respectively. The analysis of the two sets of values demonstrates the improvement of the redundancy problem.

After getting the accuracy value, we also evaluate in detail on each image produced. The correct result is compared in Fig. 6. In Fig. 6 (a)-(c), Traditional Viola Jones produces one person who has a redundant face detection, whereas in Fig. 6 (d)-(f), the proposed method produces one person who is detected correctly.

The incorrect result is compared in Fig.7. In Fig. 7 (a), Traditional Viola Jones produces one person who has a redundant face detection and one person who is not detected, whereas in Fig. 6 (d), the proposed method produces one person who is detected correctly and one person who is not detected. Meanwhile, Fig. 6 (b) and Fig. 6 (e) produce images that are equally undetectable as faces. Fig. 6 (c) and

Authorized licensed use limited to: Middlesex University. Downloaded on November 03,2020 at 16:33:53 UTC from IEEE Xplore. Restrictions apply.

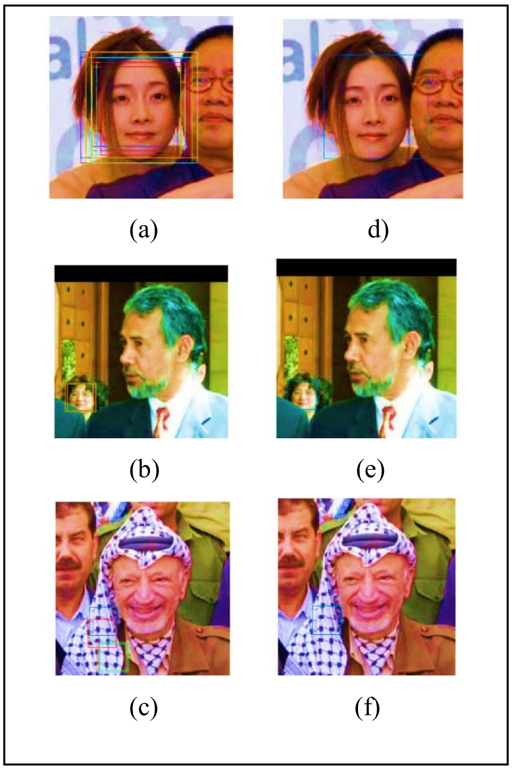
Fig. 6 (f) show face detection errors. These three pictures show that when Viola Jones cannot detect faces, the proposed method cannot detect faces as well. The conditions that cause Traditional Viola Jones cannot detect, namely: the position of the face that is tilted, the face that is not recorded



intact in the image, or the detection error produced by Viola Jones algorithms.

Fig. 6. The Correct Result (a)-(c)Traditional Viola Jones (d)-(f) The Proposed Method

Fig. 7. The Incorrect Result (a) Traditional Viola Jones (b) The Proposed Method



The correct result is compared in Fig.6 where is Fig.6(a)-(c) for Traditional Viola Jones and Fig.7(d)-(f) for the proposed method. This comparison shows the performance of the proposed method is better than traditional viola jones, both on single face and multiple face. Moreover, the Table I shows that the proposed method is able to increase the 740 detected faces that are not experience redundancy from 520 faces in the previous method. It is equivalent to repairing 220 faces (22%).

IV. CONCLUSION

In this study, we improved Viola-Jones face detection using Hill Climbing algorithm in order to reduce the redundancy rate. We have evaluated and compared the accurate, accurate and recall tests the performance of the proposed method using LFW dataset. Traditional Viola Jones achieves accuracy, precision, and recall respectively at 77%, 71%, and 81%. Meanwhile, the proposed method achieves accuracy, precision, and recall by 85%, 95%, and 88%, respectively. The analysis of the two sets of values concluded that the proposed method reduced the redundancy problem. However the problems that cannot be dealt can be addressed as the further research, such as tilted face position, faces that are not recorded intact, or face detection errors.

REFERENCES

[1] K. C. Kirana, S. Wibawanto, N. Hidayah, and G. P. Cahyono, “The Improved Artificial Neural Network based on Cosine Similarity for Facial Emotion Recognition,” *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, 2019.

[2] K. C. Kirana, S. Wibawanto, and H. W. Herwanto, “Emotion recognition using fisher face-based viola-jones algorithm,” in *International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2018, vol. 2018-Octob, doi: 10.1109/EECSI.2018.8752783.

[3] S. Wibawanto and K. C. Kirana, “Recognition of student emotion based on matrix-1 median fisher’s face and backpropagation algorithm,” in *2017 International Conference on Electrical Engineering and Informatics (ICELTICs)*, 2017, pp. 103–108, doi: 10.1109/ICELTICS.2017.8253255.

[4] Ie. of face detection accuracy of V.-J. algorithm in video applications Lopashchenko and O. Mazurkiewicz, “Estimation of face detection accuracy of Viola-Jones algorithm in video applications,” in *2018 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, 2018, pp. 1–4, doi: 10.1109/UkrMiCo43733.2018.9047552.

[5] I. G. N. M. K. Raya, A. N. Jati, and R. E. Saputra, “Analysis realization of Viola-Jones method for face detection on CCTV camera based on embedded system,” in *2017 International Conference on Robotics, Biomimetics, and Intelligent Computational Systems (Robionetics)*, 2017, pp. 1–5, doi: 10.1109/ROBIONETICS.2017.8203427.

[6] M. V. Alyushin and A. A. Lyubshov, “The Viola-Jones algorithm performance enhancement for a person’s face recognition task in the long-wave infrared radiation range,” *Proc. 2018 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. ElConRus 2018*, vol.

Authorized licensed use limited to: Middlesex University. Downloaded on November 03,2020 at 16:33:53 UTC from IEEE Xplore. Restrictions apply.

2018-Janua, pp. 1813–1816, 2018, doi: 10.1109/EIConRus.2018.8317459.

[7] S. R. Lang, M. H. Luerssen, and D. M. W. Powers, "Evolutionary feature preselection for viola-jones classifier training," *2012 Spring World Congr. Eng. Technol. SCET 2012 - Proc.*, pp. 1–4, 2012, doi: 10.1109/SCET.2012.6342142.

[8] W. Y. Lu and M. Yang, "Face detection based on viola-jones algorithm applying composite features," *Proc. - 2019 Int. Conf. Robot. Intell. Syst. ICRIS 2019*, no. 4, pp. 82–85, 2019, doi: 10.1109/ICRIS.2019.00029.

[9] H. Sharma, S. Saurav, S. Singh, A. K. Saini, and R. Saini, "Analyzing impact of image scaling algorithms on viola-jones face detection framework," *2015 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2015*, pp. 1715–1718, 2015, doi: 10.1109/ICACCI.2015.7275860.

[10] A. I. Taloba, A. A. Sewisy, and Y. A. Dawood, "Accuracy enhancement scaling factor of viola-jones using genetic algorithms," *ICENCO 2018 - 14th Int. Comput. Eng. Conf. Secur. Smart Soc.*, pp. 209–212, 2019, doi: 10.1109/ICENCO.2018.8636121.

[11] P. J. Loresco, I. Valenzuela, A. Culaba, and E. Dadios, "Viola-jones method of marker detection for scale-invariant calculation of lettuce leaf area," *2018 IEEE 10th Int. Conf. Humanoid, Nanotechnology, Inf. Technol. Commun. Control. Environ. Manag. HNICEM 2018*, pp. 1–5, 2019, doi: 10.1109/HNICEM.2018.8666244.

[12] T. Ohashi, Z. Aghbari, and A. Makinouchi, "Fast Segmentation of Texture Image Regions based on Hill-Climbing," *IEEE Pacific RIM Conf. Commun. Comput. Signal Process. - Proc.*, vol. II, pp. 848–851, 2003, doi: 10.1109/pacrim.2003.1235914.

[13] T. T. San and N. War, "Feature based disparity estimation using hill-climbing algorithm," *Proc. - 2017 15th IEEE/ACIS Int. Conf. Softw. Eng. Res. Manag. Appl. SERA 2017*, pp. 129–133, 2017, doi: 10.1109/SERA.2017.7965718.

[14] A. Pradesh, "IMPROVED HILLCLIMBING BASED SEGMENTATION (IHCBS) TECHNIQUE FOR CBIR SYSTEM."

[15] Z. Wei and J. Yan, "Automatic sulcal basins segmentation using hill climbing based on cortical surface principal curvature," *Proc. - 2011 4th Int. Conf. Biomed. Eng. Informatics, BMEI 2011*, vol. 2, pp. 653–657, 2011, doi: 10.1109/BMEI.2011.6098369.

[16] P. Ganesan, V. Kalist, and B. S. Sathish, "Histogram based hill climbing optimization for the segmentation of region of interest in satellite images," *IEEE WCTFTR 2016 - Proc. 2016 World Conf. Futur. Trends Res. Innov. Soc. Welf.*, pp. 5–9, 2016, doi: 10.1109/STARTUP.2016.7583961.

[17] P. Ganesan, B. S. Sathish, and G. Sajiv, "Automatic segmentation of fruits in CIELuv color space image using hill climbing optimization and fuzzy C-Means clustering," *IEEE WCTFTR 2016 - Proc. 2016 World Conf. Futur. Trends Res. Innov. Soc. Welf.*, pp. 3–8, 2016, doi: 10.1109/STARTUP.2016.7583960.

[18] P. Ganesan, M. Vadivel, V. G. Sivakumar, and K. Vasanth, "Hill Climbing Optimization and Fuzzy C-Means Clustering for Melanoma Skin Cancer Identification and Segmentation," *2020 6th Int. Conf. Adv. Comput. Commun. Syst. ICACCS 2020*, pp. 357–361, 2020, doi: 10.1109/ICACCS48705.2020.9074333.

[19] P. Roy, P. K. Biswas, and B. K. Das, "A Modified Hill Climbing Based Watershed Algorithm and Its Real Time FPGA Implementation," *Proc. IEEE Int. Conf. VLSI Des.*, vol. 2016-March, pp. 451–456, 2016, doi: 10.1109/VLSID.2016.103.

[20] J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool, and S. Marcel, "Hill-climbing attack to an Eigenface-based face verification system," in *2009 First IEEE International Conference on Biometrics, Identity and Security (BIdS)*, 2009, pp. 1–6, doi: 10.1109/BIDS.2009.5507530.

[21] I. Pardines, M. Sanchez-Elez, D. A. C. Martínez, and J. I. Gómez, "Online Evaluation Methodology of Laboratory Sessions in Computer Science Degrees," *IEEE Rev. Iberoam. Tecnol. del Aprendiz.*, vol. 9, no. 4, pp. 122–130, 2014, doi: 10.1109/RITA.2014.2363003.

[22] E. Learned-Miller, G. B. Huang, A. RoyChowdhury, H. Li, and G. Hua., "In Advances in Face Detection and Facial Image Analysis," *Springer*, 2016. http://vis-www.cs.umass.edu/lfw/.

[23] Luzpaz, "Haar Cascades Frontal Face," *Github*, 2013. https://github.com/opencv/opencv/tree/master/data/haarcascades.

Authorized licensed use limited to: Middlesex University. Downloaded on November 03,2020 at 16:33:53 UTC from IEEE Xplore. Restrictions apply.