Human Age and Gender Estimation using Facial Image Processing

Conference Paper · June 2020

DOI: 10.1109/TENSYMP50017.2020.9230933

2020 IEEE Region 10 Symposium (TENSYMP), 5-7 June 2020, Dhaka, Bangladesh

Human Age and Gender Estimation using Facial Image Processing

Syed Taskeen Rahman
Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
syedtrfahim@gmail.com

Asiful Arefeen
Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
asifularefeen1234@gmail.com

Shashoto Sharif Mridul
Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
shashotosharif420@gmail.com

Asir Intisar Khan
Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
asir@eee.buet.ac.bd

Samia Subrina
Dept. of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
samiasubrina@eee.buet.ac.bd

***Abstract*—The increasing demand of smart security systems has enhanced the demand for the proper identification and verification of a person. In this context, accurate estimation of age as well as proper identification of gender is highly significant. Therefore, in this work, we have implemented two separate methods with satisfactory runtime and efficiency to estimate both human age and gender using facial images. Our image processing based method involves comparison of some features extracted from the post-processed facial images of people of various age ranges followed by some edge-detection procedures, creation of binary masks and evaluation of wrinkle densities. Afterwards, thresholds were set via Naïve Bayes Classification to estimate classes. For the assessment purpose, we developed a database namely BUET facial database, which consists of images of both male and female of diverse ages. For the developed database, our proposed algorithm exhibits 76.3% accuracy in the age group classification while it shows 86.6% accuracy in the gender classification. Apart from BUET facial database, our developed algorithm has also been tested in three other databases and compared its performance with the reported literature for these databases. The mean absolute error is almost below 5.0 for this work, whereas, others exceed 5.0 in most of the cases. Moreover, the proposed algorithm exhibits reasonably good accuracy under different lighting conditions of images as well. Our study would provide further insight into the choice of appropriate features for the efficient and accurate estimation of the age and the gender of a person.**

***Index Terms*—Age and gender estimation, Computer Vision, Feature Extraction, Image Processing, Preprocessing, RGB, HSV, Cybercrime, LUT, Binary mask, Threshold, Canny Edge Detector, Gaussian Filter, wrinkles, MATLAB, Cascade Object Detector, Viola Jones Algorithm, comparison, fast, accuracy.**

I. Introduction

Ever since the development of artificial intelligence (A.I.), there have been remarkable breakthroughs in the identification of animate and inanimate objects alike [1] owing to the advancements in classification through neural networks [2] and a desire among the researchers to create smart recognition systems [3]. Recently, the use of smarter security systems involving smart devices based on biometry [4] has increased. On the other hand, the upsurge in online activity such as socializing, banking [5], shopping etc. and an influx of electronic devices and dependence on it for storing personal information have raised cybercrimes at an alarming rate [6]. A significant amount of contents on the internet is age restricted and relies upon the honesty of the user during age verification. This leads to underage citizens being able to access harmful or misleading materials [7]. An automatic age estimation, based on the image snapped on the webcam or front camera on a smart device, may prevent juveniles falling victim to online predators [8]. Additionally, the identification of gender will prevent prying offenders from creating fake accounts on social media and provide protection from cyber harassment [9]. Moreover, an automatic age and gender prediction will help social networks display materials that likely appeals to that particular viewer [5]. In order to tackle cybercrimes altogether, the cyber security personnel have started the psychological profiling of previous perpetrators and suspects, where the age and gender of the criminal is a key factor of their behavior [10].

In this context, the age and gender estimation techniques have been widely investigated. In most cases [11–13], the age has been estimated based on the common human perception of aging, where the emphasis on wrinkles is the base of estimation. The Picasso system [13] has been applied to extract the mean face whereas histogram equalization has been applied to intensify the regions with similar density values. Afterwards, an LUT [13] (Look Up Table) has been introduced to match the patterns in order to estimate the age range of 5 years interval between 15-60 years old. Another approach involved extracting features based on one independent property [11], such as color, frequency, texture and shape. Then each of those features was analyzed and the similar extracted outputs were assigned to the similar outcomes. Furthermore, age has been estimated through linear partial least squares (PLS) regression [14, 15].

Following up all these approaches, in this paper, we have implemented two separate methods with satisfactory runtime and efficiency to estimate both human age and gender. The proposed method involves careful considerations of regions of interest (ROI) and elimination of regions, therefore reducing the computational complexity. Our proposed method is tested against a database developed by ourselves namely BUET facial database and other databases as well. A comparative study is performed among the proposed method and the reported literature for each of the database.

II. Database

Our own developed database, named BUET facial database, comprises of 400 samples. There are 243 unique images of male and 157 unique images of female. The process involved facial images of a notable range of demographic. The age range starts from 16 years up to 50 years with an interspacing of 5. The camera used to capture the photos was a Nikon D5200 with a lens of 18-55 mm. The lighting was mostly sunlight, but in some cases, artificial lights such as fluorescent tubes and streetlights were used. Demonstration is done under different types of lighting conditions. We have employed MATLAB 2016a as our simulation tool.

The other databases we used are-

- *The Adience [16] database,* which consists of 26,580 images, portraying 2,284 individuals, classified for 8 age groups and gender.
- *The ORL face database,* which consists the images of 40 distinct subjects, each having 10 images in varying conditions
- *The University of Essex Face database,* which comprises of images of 225 distinct subjects, each having 20 images.

III. Methodology

For the detection of age, firstly the images are converted from RGB format to HSV. The RGB color model is a simple additive combination of the three primary colors, which represent different frequency bands in the visible spectrum, and produce different colors based on different ratios of addition. However, this may not represent the true nature of a color, as the contrast or sharpness of a color may not be achieved.

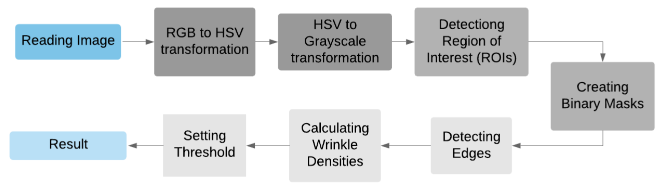


Fig. 1: Flowchart for the detection of age

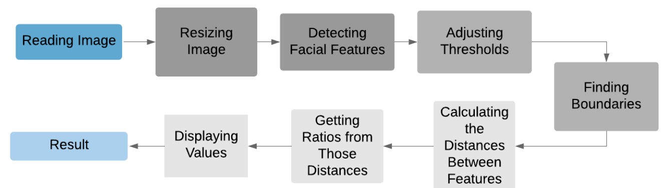


Fig. 2: Flowchart for the detection of gender

Considering this fact, we have used the HSV color model. Here, the color is given a dimension called ”Value” and the saturation or ”Brightness” of that color is also given a separate dimension [17]. Figure 3 shows a sample of RGB to HSV Transformation. Then, the HSV image was remodeled to grayscale as shown in Figure 3c in order to detect the edges i.e. wrinkles in the facial landmarks or areas of interest. Since we are looking for edges, difference in brightness perceived by “Value” is the most significant attribute.

Afterwards, three binary masks are created within the image. Coordinates of the Bounding boxes of eyes and lips calculated by the Viola Jones algorithm is used to approximate the positions of the cheeks and forehead. These are our regions of interest [18].

The mask created is a binary image of the region of interest, which is multiplied as element by element with the HSV image therefore creating an HSV mask. The edges within the binary mask are calculated utilizing the ”Canny” [19] edge detector [20, 21].

First, a Gaussian filter [22] is used to smooth the image and remove all noises. The filter function g(x,y) is given by

$$g(x,y) = \frac{1}{2\pi\sigma^2} exp^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

(a) *RGB Image* (b) *HSV Image* (c) *Grayscale Image*

Fig. 3: Transformation of the image from RGB to HSV type

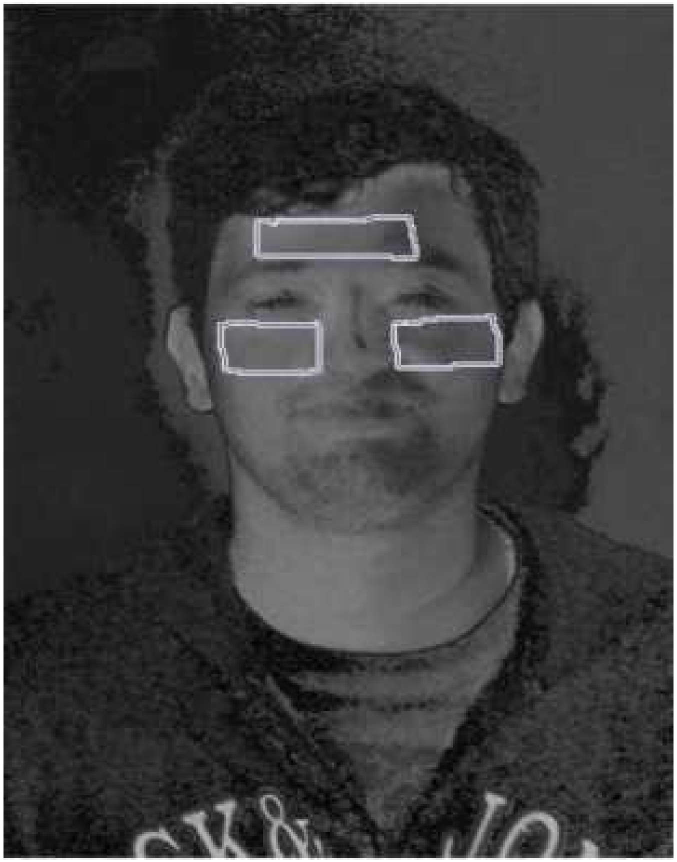


Fig. 4: Binary masks detection

where,

$g(x,y)$ = *The filter function*

σ = *The 2D standard deviation of data*

x,y = *The distances of the selected point from the origin in the horizontal and vertical axes respectively*

Equation (1) is valid for a 2D Gaussian filter with specific kernel size.

Next it finds the intensity gradient of the image which denotes the direction in which the intensity of the image pixels increases.

$$G = \sqrt{(Gx^2 + Gy^2)} \tag{2}$$

$$\theta = tan^{-1}\frac{G_y}{G_x} \tag{3}$$

where,

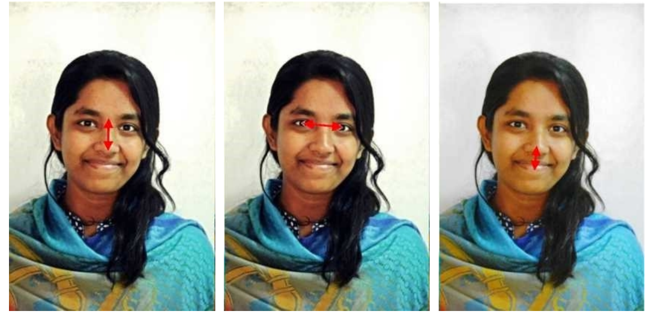
G = *Gradient of a pixel*

G_x, G_y = *The two perpendicular components of gradient*

θ = *angle at which intensity is increasing*

These two parameters use the first derivative in the vertical and horizontal direction to find the gradient of the image.

To remove unwanted points on the edges, an edge thinning algorithm is applied which sharpens the edges. It involves finding neighboring edges sharing common points and removing those common points. Finally, a threshold value is used to suppress the low strength or ”weak” edges and the remaining edges provide a better model for the real edges within the image.



(a) *Glabella to nose* (b) *Eye to eye* (c) *Lip to nose tip*

Fig. 5: Distances between various facial features

The output of a Canny edge detection algorithm applied to a binary image is a matrix which has non-zero elements, namely 1’s, in place of where it found the edges. Subsequently, the total number of pixels on the 3 matrices representing the edges are counted. A variable denoting the total number of edges or ”wrinkles” is introduced. Next, the total area of the regions of interest (ROI) is calculated. We consider the product of the number of rows and columns of the matrices as a good approximation of the total area of a region of interest and calculate the sum of these areas. The final step is to find the wrinkle density by dividing the number of edges by the total area and the output is employed as the threshold value for distinct ranges of age. The threshold is created by feeding a training set in a Naïve Bayes Classifier where the classifiers are the different ranges of age.

For the determination of gender, the image is resized to a standard size of *512*512* pixels since the relative positions of facial features are to be measured. Next, the dynamism of the image processing and computer vision toolboxes of MATLAB is utilized to create an object of the class vision. CascadeObjectDetector [23], using the ViolaJones algorithm [23, 24], is employed to find certain facial landmarks. Here, we create four objects to find landmarks, namely, the mouth, the lips and the two eyes. Afterwards, a boundary is created for all four landmarks, each having a center to be based upon. The step function creates an *N*4* matrix, where N is the number of objects under a certain class, and each class has 4 properties; height, width and the co-ordinates of the center. The program is designed to iterate until there are precise threshold values so that one object may only have one occurrence, because there can only be one nose or one left eye. Thereafter, the ratios between the relative sizes and distances between the landmarks are calculated. The calculations include the distance between the two eyes d_1, the distance between lip to nose tip d_2 and the distance between the nose and the apparent center between eyes d_3. These give two ratios to be utilized, the ratios d_1/d_2 and d_3/d_2. Finally, the sum of these ratios spells a value referring to the possible gender of a person.The threshold value employed is calculated by feeding a training set in a logistic regression algorithm.

IV. Result & Performance Analysis

In order to investigate our proposition, the developed algorithm is experimented with both our own database and other databases also. The images is subjected to various lighting and background conditions. Table I displays a part of our (BUET facial database) database, holding 30 distinct images, along with their predicted and actual ages and genders.

A comparative study on accuracy between our developed method and a method proposed by Eidinger *et al.* [16] is presented in Table II. Both methods employ the Adience [16] database for benchmarking during experimentation. It is to be mentioned that age group of BUET facial database has been altered to match those of our counterpart who employed a more compact age group classification.

Based on BUET facial database and the other considered benchmarked databases, we have calculated the mean value of runtime of the program for all images and it is 4.17 seconds. There was no significant difference in runtimes for different gender or age group combinations, whether the estimation is correct or incorrect. For all the correct cases, the runtime average was 4.12 seconds and for incorrect cases such as age and/or gender, the value was 4.63 seconds. Table III and Table IV give a comparative analysis on accuracy to estimate age and gender using different methods. Table IV depicts that on an average the proposed method give 5.43% and 9% higher accuracy in comparison to other methods.

While developing our method, we concede that lighting conditions played a significant role in the accuracy of the outcomes. Facial Images open for public use were tested upon produced more erroneous results, owing to the fact that they had diverse lighting conditions and imaging sources.

Here in Table IV, ***No. of subjects*** refers to the number of distinct persons we took photos of and ***Total Images*** indicates the accumulated number of all photos.

V. Conclusion

In summary, an efficient and feasible algorithm has been proposed and implemented in this paper. It is computationally fast and auspicious as the algorithm developed in this work is not serpentine and consists of edge detection, creation of binary masks and estimation of wrinkle densities. It performs with good accuracy on all types of databases. The proposed algorithm is tested on a number of images under different conditions with an accuracy of 58.7% (for age) and 82% (for gender). This execution time is quite reasonable and good enough. Furthermore, for the same benchmark, this procedure produces better result than the methodology claimed by Eidinger *et al.* Accuracy of similar level is both challenging to achieve and quite unique compared to Eidinger *et al.* Quite undeniably, the programme developed here is performing better than similar works produced till date. This work has copious potential and numerous applications. Such study would encourage further investigation for the choice of appropriate features for the estimation of human age and gender.

TABLE I: A few samples and results from BUET facial database

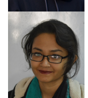
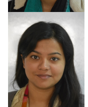
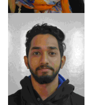
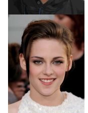
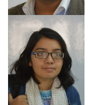
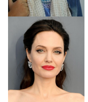
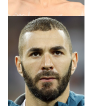
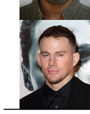
Specimen	Predicted age	Actual age	Predicted gender	Actual gender
	21-25	21	Male	Male
	21-25	21	Male	Male
	21-25	22	Male	Male
	21-25	21	Male	Male
	31-35	22	Male	Male
	21-25	21	Female	Female
	26-30	23	Female	Female
	21-25	22	Female	Female
	21-25	22	Male	Male
	16-20	28	Female	Female
	21-25	22	Male	Male
	21-25	23	Female	Female
	41-45	43	Male	Female
	26-30	29	Female	Male
	21-25	21	Male	Male
	21-25	22	Male	Male
	51-55	38	Male	Male

TABLE II: Performance(accuracy) comparison between the developed method and a method proposed by Eidinger [16] against the same benchmark

Proposed Method	Eidinger *et al.* [16]
Age Group: 58.7% (Best score)	Age Group: 45.1%
Gender estimation: 82%	Gender Estimation: 88.6% (Best score)

TABLE III: Comparison of accuracy among different methods of age estimation

Work	Database	MAE
This work	BUET facial database	4.6
	ORL database	4.3
	Adience	4.9
	University Of Essex	5.1

MAE : Mean Absolute Error between real age and estimated age

Publication	Database	Accuracy (age)
Lanitis *et al.* [25]	Private	MAE : 4.3
Hayashi *et al.* [13]	Private	Age: 27%
Choi *et al.* [26]	FG-NET, PAL, Private	Agg: 68%
Fu and Huang *et al.* [27]	Private	MAE : 6
Guo *et al.* [14]	FG-NET, Private	MAE : 4.8
Guo and Wang *et al.* [12]	UIUC,FG-NET	MAE : 5.7
Thukral *et al.* [10]	FG-NET	MAE : 6.2

Agg: Aggregated/Cumulative accuracy

VI. Acknowledgment

The authors acknowledge the contribution of every random person who participated in forming the BUET Facial Database. The authors would like to thank the Robert Noyce Laboratory of Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology (BUET) for providing necessary experimental platform.

References

[1] P.F. Felzenszwalb and D.P. Huttenlocher, "Pictorial Structures for Object Recognition," *International Journal of Computer Vision*, vol. 61, issue. 1, pp. 55-79, January 2005.

[2] Stephan Dreiseitl and Lucila Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *Journal of Biomedical Informatics*, vol. 35, issue. 5-6, pp. 352-359, October 2002.

[3] Alex Pentland and Tanzeem Choudhury, "Face Recognition for Smart Environments," *IEEE Computer*, vol. 33, pp. 50-55, 2000.

[4] Khitrov, Mikhail, "Talking passwords: Voice biometrics for data access and security," *Biometric Technology Today*, vol. 13, pp. 9-11, February 2013.

[5] B. Vatanasombut, M. Igbaria, A. Stylianou and W. Rodgers, "Information systems continuance intention of web-based applications customers: The case of online banking," *Information & Management*, vol. 45, issue. 1, pp. 419-428, November 2008.

[6] Katharina Krombholz, Dieter Merkl and Edgar Weippl, "Fake identities in social media: A case study on the sustainability of the Facebook business model," *Journal of Service Science Research*, vol. 4, issue. 2, pp. 175-212, December 2012.

[7] Robert Slonje, Peter K. Smith, Ann Frisén, "The nature of cyberbullying, and strategies for prevention," *Computers in Human Behavior*, vol. 29, issue. 1, pp. 26-32, January 2013.

[8] M.L. Ybarra and K.J. Mitchell, "How risky are social networking sites? A comparison of places online where youth sexual solicitation and harassment occurs," *PubMed*, vol. 121, February 2008.

[9] A. Sengupta and A. Chaudhuri, "Are Social Networking Sites a Source of Online Harassment for Teens? Evidence from Survey Data," *Children and Youth Services Review*, vol. 33, pp. 284-290, February 2011.

[10] P. Thukral, K. Mitra, and R. Chellappa, "A hierarchical ap-. proach for human age estimation," *IEEE ICASSP*, pp. 1529 –1532, 2012.

[11] H. Takimoto, T. Kuwano, Y. Mitsukura, H. Fukai and M. Fukumi, "Appearanceage feature extration from facial image based on age perception," *SICE Annual Conference*, pp. 2813-2818, 2007.

[12] G. Guo, Y. Fu, C. R. Dyer and T. S. Huang, "Image-Based Human Age Estimation by Manifold Learning and Locally Adjusted Robust Regression," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1178-1188, July 2008.

[13] J. Hayashi, M. Yasumoto, H. Ito and H. Koshimizu, "Age and gender estimation based on wrinkle texture and color of facial images," *Object recognition supported by user interaction for service robots*, vol. 1, pp. 405-408, August 2002.

[14] G. Guo and G. Mu, "Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression," *IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 657-664, 2011.

[15] W. R. Schwartz and L. S. Davis, "Learning Discriminative Appearance-Based Models Using Partial Least Squares," *2009 XXII Brazilian Symposium on Computer Graphics and Image Processing*, pp. 322-329, October 2009.

[16] Eran Eidinger, Roee Enbar, Tal Hassner, "Age and Gender Estimation of Unfiltered Faces," *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*, vol. 9, pp. 2170-2179, 2013.

[17] S. Sural, G. Qian and S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval," *Proc. Int. Conf. Image Process.*, vol. 2, pp. 589, February 2002.

[18] Y. Wu, .P Kalra, L. Moccozet, N. Magnenat-Thalmann, "Simulating wrinkles and skin aging," *The visual computer*, vol. 15, pp. 183-198,1999.

[19] J. Canny, "A Computational Approach To Edge Detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, pp. 679 - 698, December 1986.

[20] K. Yow and R. Cipolla, "Feature-based human face detection," *Image Vis. Comput.*, vol. 15, pp. 713-735, September 1997.

TABLE IV: Comparison of accuracy to estimate age and gender using different methods

Database	***No. of subjects***	***Total Images***	***Accuracy in Age Group Classification***	***Accuracy in Gender Classification***
BUET facial database	150	150	76.3%	86.6%
ORL	40	400	80.5%	72.7%
Adience	2284	26580	67.8%	82%
University of Essex	225	4500	64.3%	78.1%

[21] K. Ghosh, Swapan Sen, S. Sarkar, R.K. Saha and S. Karmakar, “A new design of low pass filter by Gaussian derivative family,” *Third International Conference on Intelligent Sensing and Information Processing*, vol. 1, pp. 177-182, 2005.

[22] Deriche, R., “Using Canny’s criteria to derive a recursively implemented optimal edge detector,” *International Journal of Computer Vision*, vol. 1, issue. 2, pp. 167-187, June 1987.

[23] P. Viola and M. Jones, “Fast and robust classification using asymmetric adaboost and a detector cascade,” *NIPS*, 2001.

[24] P. Viola, and M. Jones, “Robust real-time face detection,” *International Journal of Computer Vision*, vol. 57, pp. 137-154, 2004.

[25] A. Lanitis, C. Draganova, and C. Christodoulou, “Comparing different classifiers for automatic age estimation,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, issue. 1, pp. 621-628, February 2004.

[26] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park and J. Kim, “Age estimation using a hierarchical classifier based on global and local facial features,” *Pattern Recognition*, vol. 44, issue. 1, pp. 1262-1281, June 2011.

[27] Y. Fu and T. S. Huang, “Human Age Estimation with Regression on Discriminative Aging Manifold,” *IEEE Transactions on Multimedia*, vol. 10, no. 4, pp. 578-584, 2008.