Computer Vision-Based Gender Detection from Facial Image

Emon Kumar Dey
University of Dhaka

Ali Md. Haider
BRAC University

Computer Vision-Based Gender Detection from Facial Image

Emon Kumar Dey, Mohsin Khan, & Md Haider Ali

Manuscript
Received:
10,May, 2013
Revised:
25,May, 2013
Accepted:
9, Jul., 2013
Published:
15, Jul., 2013

Keywords
Gender detection, face detection, object detection, image processing

***Abstract*— Computer vision-based gender detection from facial images is a challenging and important task for computer vision-based researchers. The automatic gender detection from face images has potential applications in visual surveillance and human-computer interaction systems (HCI). Human faces provide important visual information for gender perception. This research presents a novel approach for gender detection from facial images. The system can automatically detect face from input images and the detected facial area is taken as region of interest (ROI). Discrete Cosine Transformation (DCT) of that ROI plays an important role in gender detection. A gender knowledgebase of the processed DCT is created utilizing supervised learning. To detect gender, input image is passed through a classifier which is based on that knowledgebase. To improve the matching accuracy Local Binary Pattern (LBP) of the ROI is done before converting it into DCT. This research has experimented on a database of more than 4000 facial images which are mainly of this subcontinent in order to evaluate the performance of the proposed system. The average accuracy rate achieved by the system is more than 78%.**

1

1. Introduction

Recently gender classification from face image is an attractive research topic and one of the actual problems of computer vision. If a person prevent from being seen his hairstyle, remove his facial hair and makes other changes to his face, human can still detect the sex with an accuracy of more than 90%. This observation attracts most of the scientist that what is the facial information by which human can make difference between men and women. A human can detect gender from face easily but it is very challenging task for computer which has no intelligence. In modern world everything is going to be machine dependent. With the growing demand for security, reliability, convenience, computer vision approaches such as face detection, gesture detection, person identification, motion detection and perhaps most fundamentally gender detection will play important role in our life. Target of this research is to develop an intelligent system for identifying the gender of a person by the facial image. Sectors such as biometric authentication, high-tech surveillance, security systems, criminology, automatic psycho physiologic inspection, etc. can be benefited by the automatic gender detection system.

There are mainly two approaches. The first one came from the psychophysical explorations of human face. It generally used the features such as distances between nose, eyes and mouth, areas of different face parts and so on. The second approach use the low level information of the face image which is based on the image pixels of the face. The first approach also known as the high level features based classification. But this first method has some problem for automatic detection. The result is not very accurate and unsolved in many cases. This research is mainly focused on the second approach which is based on low level information approach. This approach examines the pixel properties of different area of a face image and makes decision about the gender of the face image.

2. Related Works

The sex of a face is conveyed by several classes of information, namely (a) local features (such as facial hair, eyebrows, and skin texture), (b) configural relationships between features, and (c) the 3D structure of the face. The shape differences between the two sexes, and compared the average male and average female heads which were obtained using a laser range scanner. On average, the male face has a more protuberant nose, brow, chin/jaw than the female face. The female face, on the other hand, has somewhat more protrusive cheeks than the male face. Moreover, the greatest differences were found in the regions of the nose and chin. Cottrell et al [1] proposed a method which reduced the dimension of whole face images by applying auto encoder network and classified gender based on the reduced input features. Tamura et al [2] used a neural network and proved that if the image is even very low resolution such as 8 x 8 can be used for gender classification. Jain and Huang [3] extracted the features by an approach which known as independent component analysis (ICA) and classified it with linear discriminate analysis (LDA). Burton et al.[4] extracted point-to-point distances from 73 fixed points on face images and used discriminant analysis as a classifier. Brunelli and Poggio (1992) extracted 16 geometric features such as eyebrow thickness and pupilto- eyebrow distance and used HyperBF networks as a classifier. BenAbdelkader and Griffin [5] proposed a method. This method extracted regions from the face. The region is then used for the input of

This work is supported by University of Dhaka, Bangladesh.
1, ✉ *Emon Kumar Dey is with ...(Email: emonkd@unívdhaka.edu),*
2, ✉ *Mohsin Khan is with ...(Email: mohsíncsedu@yahoo.com),*
3, ✉ *& Md Haider Ali is with ...(Email: haíder@unívdhaka.edu).*

SVM or Fisher Linear Discriminant (FLD) gender classifier. The eye locations were used to align the face. For the experiments, they used images obtained from various databases. Hyun-Chul Kim et al.[6] proposed an appearance based gender classification method. They concern the gender classification task of discriminating between images of faces of men and women from face images. The initial images are preprocessed and normalized and input into classifiers. They propose to use Gaussian process classifiers which are Bayesian kernel classifiers. Wiskott et al. [7] presented a system where a model graph was placed manually to a face and gender classification was based on the Gabor wavelets placed on the model nodes. They also used the system for face recognition.

3. Proposed Gender Detection Method

In the proposed gender detection method the first step is to detect facial area from an input image. We use Viola and Jones [8] method to detect facial region from an input image. The detected face area is then passed through a classifier which identifies the gender of the detected face by analyzing some properties and attributes of the face image. Figure 1 shows the proposed gender detection system. Each part of the work will be described in the following subsections. This method mainly experimented on Bangladeshi facial images. The database is collected from Dhaka University admission office. In this type of research ethnicity plays a very important role [16].

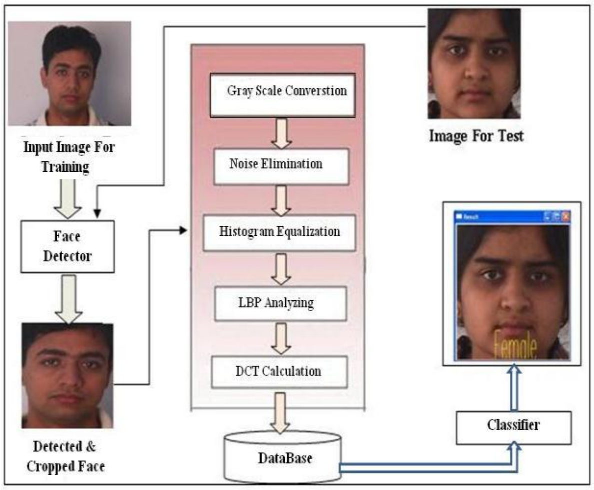


Fig. 1. Proposed Gender Detection System from Face Image.

A. Face Detection

The viola and Jones face detection algorithm follow the basic principle that a sub window is scanned for finding and detecting the face from an image. This approach to detecting objects or face in images combines four key concepts. They are: 1)Simple rectangular features called Haar Features, 2)Integral Image for rapid face detection, 3)The Adaboost machine learning method ,4)A Cascaded Classifier to combine many features efficiently.

B. Noise Filtering and Histogram Equalization

After detecting face area from an input image a ROI of 128 x128 is selected and cropped. This size is selected because 128 x 128 image will passed through a DCT translator. This DCT subdivide the whole image into several 8 x 8 sub-region. After facial area detection, cropping and resizing the image is passed through a noise filter.

Fig.2. Face Detection Using Viola and Jones Method.

This research adopts a simple technique for making calculation easier. It is assumed that all faces resides on centered, so a bounding region can be constructed into centered on the image and consider only the image contents within this region. Any part outside the region is considered as noise and eliminated.

Histogram equalization is a technique which consists of adjusting the gray scale of the image so that the gray level histogram of the input image is mapped onto a uniform histogram. Histogram equalization is applied on the detected face image which was converted into gray scale and resized. Simple algorithm used for histogram equalization is given:

- For every pixel in the image get gray value in variable ***i***, *hist[i]* = *hist[i]* +1 when ***i*** = 0 to *255*.
- From *hist[ ]* array find the cumulative distribution frequency (cdf).
  $hist_{(cdf)}[i] = hist_{(cdf)}(i-1) + hist[i]$
- Calculate Equalized Histogram (EH) for each new *i* by following equation.

$$EH[i] = \left[\frac{(255 * hist_{(cdf)}[i]) - MxN}{MxN}\right]$$

Where, *MxN* is size of the input image

C. Local Binary Pattern (LBP) Analyzing for Feature Extraction and Discrete Cosine Transformation (DCT) of the Analyzed Image.

The original LBP operator, introduced by Ojala et al. [9]. LBP operator was originally defined to characterize textures of an image. A binary number or its equivalent decimal number is uses to characterize each pixel of the image. In the most basic version a 3 x 3 neighborhood around each pixel is considered to obtain the information about this specific pixel.

The neighbors which value is brighter than or equal the center is given a binary value 1 and those are less bright then the central pixel are given a value 0. A 3x3 area cannot capture a large enough area structure to dominate features of some textures by calculating these features. For this reason the LBP operator extended to use different neighborhood size.

Fig. 3. (a) Main Image (b) Gray Image

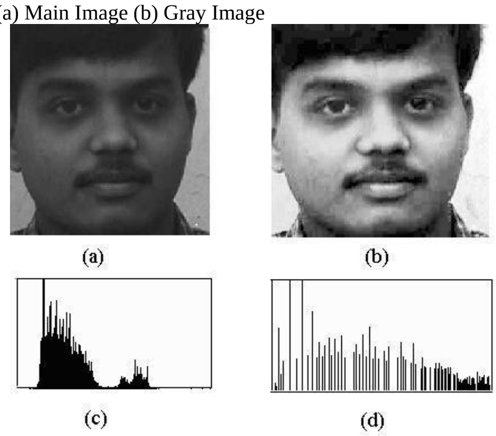


Fig. 4. (a) gray image, (b)Histogram equalized Image, (c),(d) are corresponding histogram of (a) and (b)

Where the thresholding function s(x) is defined as:

$$S(x) = \begin{cases} 1 \; if \; (lx) \geq mid \; value \\ 0 \; if \; (lx) < mid \; value \end{cases}$$

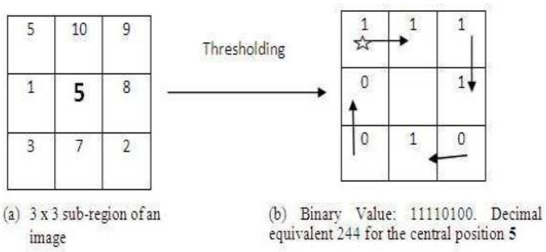


Fig. 5. LBP operator (a) a 3x3 sub region of an Image. (b) After thresholding.

Figure 6 show the image after histogram equalization and LBP operation.

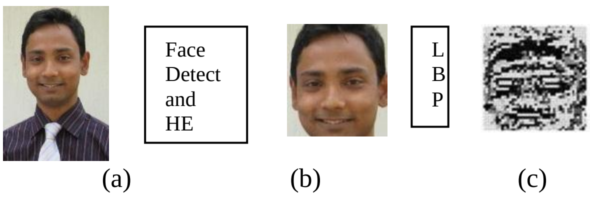


Fig. 6. LBP operation on an Image (a) Main image (b) histogram equalized (128x128) sized image (c) Image after LBP operation.

The image which is obtained after LBP operation is passed through DCT module. DCT was mainly developed to compression of image which represents an image as a sum of sinusoids of varying magnitudes and frequencies. The main interesting point of DCT is that, for a typical image most of the visually significant information about the image is concentrated in just a few co-efficient of the DCT. The DCT is close relative to Discrete Fourier Transformation (DFT). When the LBP image is passed through DCT then the visually significance information or points of the image will be in the top position of a two dimensional array of specific block. So only by examining the top points of a DCT converted image the properties of the image can be tested. These following equations are used for DCT calculation as well as transformation of a two dimensional image.

$$D(u,v) = \frac{2}{\sqrt{M*N}} a(u)a(v)x \sum_{M=0}^{M-1} \sum_{N=0}^{N-1} I(m,n) Cos[A] Cos\,[B] \quad \text{(Equ. 1)}$$

Where,

$$A = \frac{(2m+1)u\pi}{2M}, \; B = \frac{(2n+1)v\pi}{2N}$$

$$a(u) = \begin{cases} \sqrt{1/M}, & for\ u = 0 \\ \sqrt{2/M}, & for\ u = 1,2,3,\ldots\ldots\ldots M-1 \end{cases}$$

$$a(v) = \begin{cases} \sqrt{1/M}, & for\ v = 0 \\ \sqrt{2/M}, & for\ v = 1,2,3,\ldots\ldots\ldots N-1 \end{cases}$$

Here D(*u*, *v*) is the points of transformed image and I(*m*, *n*) is the image which is obtained from the main image after performing LBP operation. The whole image 128x128 is divided into several block of 8x8. DCT is performed on each 8x8 blocks. In each block equation (3.1) is applied. This values are then stored in a two dimensional array and then transformed in a one dimensional array by scanning the whole 8 x 8 array in a zigzag order. In this fashion all of the 8x8 blocks in an image are stored in a one dimensional array. This is similar to sorting the values according to importance of the pixels. High importance coefficients are located on the top-left corner of the block. So if the whole image of size 128 x128 is divided into several 8 x 8 blocks then there are total 256 blocks. So if the most significant coefficient from each block is selected which is situated on the top left corner, then total 256 points will be selected.

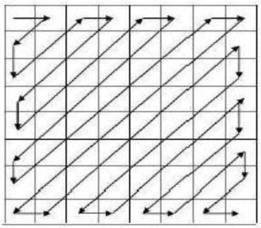
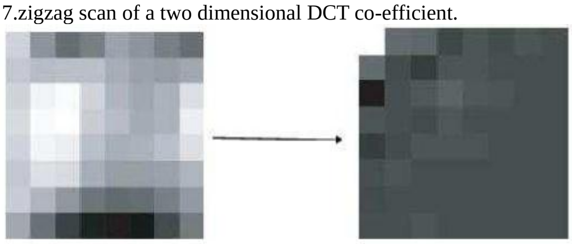


Fig. 7.zigzag scan of a two dimensional DCT co-efficient.



(a) (b)

Fig. 8. DCT transformations (a) an 8x8 image block, (b) transformed DCT of the 8x8 Block

(a) **(b)** **(c)**

Fig. 9. DCT of an Image (a) Original Image, (b) LBP of original image, (c) DCT of LBP image.

Here we can see that in the transformed figure of 8(b) most significant value of the image is in the top left position of the image (b). Figure 9 shows the original DCT output of a test image.

D. Gender Detection Classifier

This classifier is an algorithm that stores all available examples and classifies new instances of the example based on a similarity measure. There is nearly information of 3000 female faces and 3000 male faces are stored in this database. After DCT the first 4 top most values of each module of each image are then stored in a one dimensional array. So for a 128 x 128 image only 1024 pixel values are stored in a file for each of the training image (128 x 128 image contains total 256 blocks of size 8 x 8. So if we take every top most values from each block we get 256*4=1024 values). The classifier then calculates a manhattan distance between every point from this image to the training images. If the average distance from the test image to a training image is greater than or less a predefined threshold value then a hit score or miss score for the test image is increases. The hit score and miss score indicate the result for male or female. If the hit score for male is greater than the female then the test image is considered as a male face. Beside this if the hit score for female is greater than the male then the test image is considered as a female face. The detail of this technique is as follow:

- **Step 1:** Read the test image points.
- **Step 2:** For every training image calculate the distance between every point of the test image and every points of the template images which is obtained after DCT operation.
- **Step 3:** Increase the hit score by 1 for specific gender If, $Tst(x, y) - Tmp(x, y) >= threshold$,else increase miss score
- **Step 4:** Find the total hit score for the specific gender. If hit score for male is greater than the hit score for female then the test image is considered as male and vice versa

Figure 10 shows the final diagram of the proposed gender detection system.

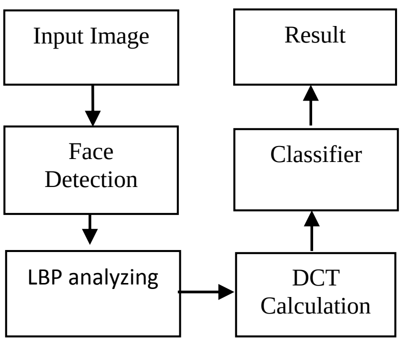


Fig. 10. Steps of Proposed Gender Detection System.

4. Experimental Result of the Proposed Gender Detection System:

This section will present experimental result with discussion. A database of trained image is stored. Information of each of these images is stored after face detection, normalization LBP and DCT translation. An experimental result of the proposed system shows that without histogram equalization the accuracy of the system is approximately less than 70% and after histogram equalization the performance significantly increases up to 80%.

The system is tested for various conditions of the input images. Table 4.1 shows the performance evaluation of the proposed gender detection method for the image in which female having long hair which is visible on the facial image.

TABLE 1:
PERFORMANCE OF DETECTING FEMALE WITH LONG HAIR VISIBLE ON THE FACIAL PORTION.

Sequence	No. of Images(T)	Detected (C)	False Detection	Accuracy A_p (%)
#1	1000	872	128	87.2
#2	1000	908	92	90.2
#3	1000	779	221	77.9
#4	1000	805	195	80.5
#5	1000	774	226	77.4
				Average = 82.64%

Here, Accuracy

$$A_p = \frac{C}{T} * 100 \qquad \text{(Equ. 2)}$$

A_p = Accuracy of the system.
C= Number of correct detection
T= Total input.

The following table will show the performance of the proposed system for detecting the female faces having short hair. Table 2 shows the result. This is tested for 764 different female faces having short hair.

TABLE 2
ACCURACY FOR DETECTING FEMALE WITH SHORT HAIR

Sequence	No. of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	150	87	63	58.00
#2	135	97	38	71.85
#3	99	65	34	65.65
#4	170	112	58	65.88
#5	210	156	54	74.28
				Average=67.13%

Table 3 will show the performance of the proposed system for detecting the female having ornaments on ears and forehead. This is tested for 1762 different female faces having ornaments on ears and forehead.

TABLE 3
ACCURACY FOR DETECTING FEMALE HAVING ORNAMENTS ON EAR AND FOREHEAD.

Sequence	No. of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	475	398	77	83.78
#2	380	314	66	82.63
#3	200	182	18	91.00
#4	370	341	29	92.16
#5	337	301	36	89.31
				Average =87.78%

Next table will show the performance of the proposed system for detecting the male having no hair (totally bald male). Table 4 shows the result. This is tested for 70 male faces of bald head.

TABLE 4
ACCURACY FOR DETECTING MALE HAVING NO HAIR.

Sequence	Number of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	10	10	0	100
#2	16	14	2	87.5
#3	10	10	0	100
#4	16	16	0	100
#5	18	16	2	88.89
				Average = 95.27%

The performance of the proposed system for detecting the male having bearded and mustache in the face is showed on table 5 and detecting male faces from input images which have a cleaned shaved shown on table 6. These are tested for 205 and 600 different male faces respectively.

TABLE 5
ACCURACY OF THE SYSTEM FOR DETECTING MALE HAVING BEARDED AND MUSTACHE.

Sequence	Number of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	55	54	1	98
#2	40	38	2	95
#3	25	25	0	100
#4	40	38	0	95
#5	45	39	1	97.5
				Average = 97.1%

The system shows its worse performance in the case when input image of a male contain long hair with ornaments and a female with short hair with no ornaments on the face. Result is shown on Table 7

TABLE 6
ACCURACY OF THE PROPOSED SYSTEM FOR DETECTING MALE HAVING CLEANED SHAVED.

Sequence	Number of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	120	100	20	83.33
#2	100	74	26	74.00
#3	100	82	18	82.00
#4	160	104	56	65.00
#5	120	76	24	64.00
				Average = 73.67%

TABLE 7
WORST PERFORMANCE OF THE SYSTEM.

Sequ-ence	Number of Images	Detected (C)	False Detection	Accuracy A_p (%)
#1	30 (male with long hair &	16	14	53.33
#2	120 (Female with short hair & no	53	67	44.16
				Average = 48.78 41%

So by considering table from 1 to 7 the final average accuracy of the system is 78.91%.

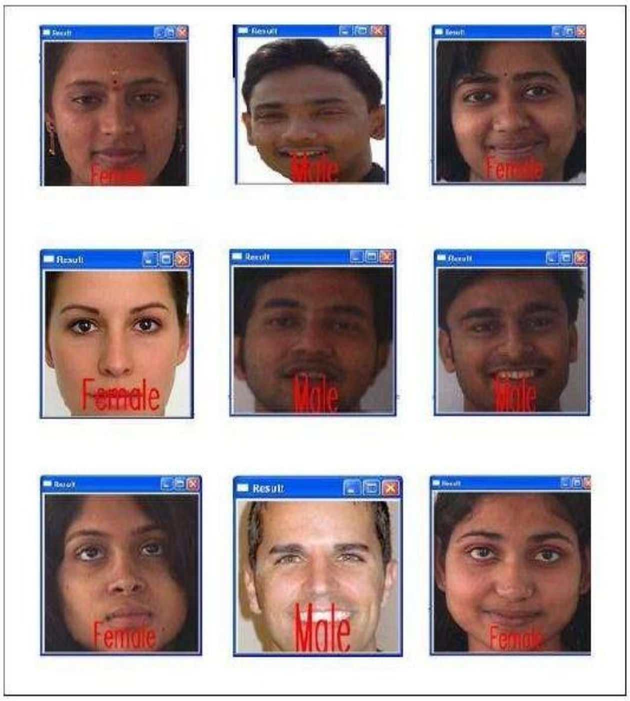


Fig. 12. Some Sample Output

5. Conclusion

This work describes the gender detection from a facial image in a novel approach with satisfactory success rates. The advantage of this system is that it takes very small amount of time to detect gender from a face image. It takes less than 20 milliseconds in a 128 X 128 pixel image on a 1.8 GHz core 2 duo processor to identify a specific gender. Average accuracy of the system is 78.91%. From the experimental result we also found that performance degrades for those facial images which contain glasses on their eyes. So for better performance we have to provide these system facial images without glasses on eyes. This system does not work properly when the face images is not in front view. Angles faces reduce the performance. So our future goal is to work on the faces with glasses and angle face image.

References

[1] Representing face images for emotion classification, Cottrell and Metcalfe., 1997 , Citeseer

[2] Male/female identification from 8 X 6 very low resolution face images by neural network, Tamura, S. and Kawai, H. and Mitsumoto, H., Pattern Recognition, 29, 2, 331–335, 1996, Elsevier

[3] Integrating independent components and linear discriminant analysis for gender classification, Jain, A. and Huang, J., Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on, 159–163, 2004, IEEE

[4] What’s the difference between men and women? Evidence from facial measurement,Burton, A.M. and Bruce, V. and Dench, N., PERCEPTION-LONDON-, 22, 153–153,1993, PION LTD

[5] Comparing and combining depth and texture cues for face recognition, BenAbdelkader, C. and Griffin, P.A., Image and Vision Computing, 23, 3, 339–352, 2005, Elsevier

[6] Appearance-based gender classification with Gaussian processes, Kim, H.C. and Kim, D.and Ghahramani, Z. and Bang, S.Y., Pattern Recognition Letters, 27, 6, 618–626,

[7] Face recognition by elastic bunch graph matching, Wiskott, L. and Fellous, J.M. and Kuiger, N. and von der Malsburg, C., Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19, 7, 775–779, 1997, IEEE.

[8] A unified learning framework for real time face detection and classification, Shakhnarovich, G. and Viola, P.A. and Moghaddam, B., Automatic Face and Gesture Recognition, 2002.Proceedings. Fifth IEEE International Conference on, 14–21, 2002, IEEE

[9] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *TPAMI*, vol. 24, pp. 971-987, 2002.

[10] Robust real-time face detection, Viola, P. and Jones, M.J., International journal of computer vision, 57, 2, 137–154, 2004, Springer

[11] An experimental comparison of gender classification methods, M”akinen, E. and Raisamo, R., Pattern Recogn. Lett. 29, 10, 1544-1556, 2008, Citeseer

[12] Gender Classification Based on 3D Face Geometry Features Using SVM, Xia Han, HassanUgail, Ian Palmer, Cyberworlds, International Conference on,, 114-118, 2009,

[13] Automatic Gender Recognition Using Fusion of Facial Strips, Ping-Han Lee, Jui-Yu Hung, Yi-Ping Hung. International Conference on Pattern recognition, IEEE 2010.

[14] The role of face parts in gender recognition, Andreu, Y. and Mollineda, R., Image Analysis and Recognition, 945–954, 2008, Springer

[15] Gallagher, A.C., Chen, T., 2009. Understanding images of groups of people,
in: Proc. of Int. Conference on Computer Vision and Pattern Recognition,
pp. 256-263.

[16] Guodong Guo; Guowang Mu, "A study of large-scale ethnicity estimation with gender and age variations," *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on* , vol., no., pp.79,86, 13-18 June 2010

[17] Kun Duan; Parikh, D.; Crandall, D.; Grauman, K., "Discovering localized attributes for fine-grained recognition," *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* , vol., no., pp.3474,3481, 16-21 June 2012

[18] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. PAMI, 33:898–916, 2011

[19] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In CVPR, 2003.

Emon Kumar Dey
Lecturer
Institute of Information Technology
University of Dhaka, Dhaka 1000.

Mohsin Khan
Lecturer
Department of Computer Science & Engineering
University of Dhaka, Dhaka 1000.

Dr. Md. Haider Ali
Professor
Department of Computer Science & Engineering
University of Dhaka, Dhaka 1000.