

# The Effect of Gender-Specific Facial Expressions on Face Recognition System's Reliability

Mohammadreza Azimi

Institute of Electronics and Information Technology,  
Warsaw University of Technology,  
ul. Nowowiejska 15/19, 00-665 Warsaw, Poland  
M\_r\_azimi1991@yahoo.com

Andrzej Pacut

Biometric Laboratory  
Research and Academic Computer Network (NASK)  
Wawozowa 18, Warsaw, Poland.  
Institute of Control and Computation Engineering  
Warsaw University of Technology  
Nowowiejska 15/19, Warsaw, Poland.

**Abstract**— previous studies concerned with Gender differences in psychology, reported that women are more emotional in comparison with men. However, from biometric science point view this fact has not been investigated yet. In this study, we have used Psychological Image Collection at Stirling (PICS) <http://pics.stir.ac.uk/>, contains pictures from 13 women and 10 male users with different emotional moods, in order to test if, as the psychologists say, women show more intense facial expression than men or not? For this purpose, two state of the art codes, Python face recognition dlib and VeriLook neurotechnology have been implemented.

**Keywords**— Gender Dependency; Facial Recognition System; Mood Expressions; Performance Analysis; Biometrics.

## I. INTRODUCTION

The human face can be described as a window into the emotional experiences of a user and facial emotion expressions are reactions which include many interconnecting elements of movements of facial muscles.

The expressions of face can give us this ability to discriminate the emotional situations from each other.

Facial expressions can be generally classified in seven different groups of fundamental emotions: Anger, disgust, fear, happiness, sadness, and surprise along with the neutral state.

In different societies with different cultural rules, by investigating the role of user gender in her/his daily social activities, it has been proved that women take part more often in family affairs that need rapid receptiveness and strong responsiveness to emotion. Based on the mentioned fact, it can be hypothesized that women show their feelings in more intense way and the facial expressions must be more intense in women than men.

As a consequence it can be accepted that verification of female users with facial expression can challenge the reliability of biometric system more.

Nowadays, due to the strengths of mobile biometric authentication and identification in comparison with traditional methods such as pattern, pins and password, and due the fact that the number of emerging options for this

purpose on advanced smartphones as they become more smarter, the use of such technology is rapidly growing. On the other hand, the urgent need for more reliable mobile biometric systems is increasing as the popularity of such systems enhanced especially between young adults.

The sufficient data is essential for the algorithm of extraction in order to recognize and categorize identities.

Additionally, we need a database which meets the mentioned requirements. However there are many social issues that affect the biometric technology. User-related parameters which include physiological factors, such as age, and behavioral factors such as habituation, cultural restrictions can influence the biometric sample characteristic.

There are a lot of social problems and human factors (Like emotional mood of participants) which can affect the reliability of biometric recognition system. However, it is still an open problem, if the gender of user can be considered as key factor which can intensify the influence of the mentioned parameters on reliability of system of biometric recognition?

In this paper, firstly in section 2, the psychological related articles concerned with gender differences in facial mood expression will be briefly reviewed, in section 3 the methods and materials will be explained and the results and discussions and conclusions parts will be presented in section 4 and 5 respectively.

## II. RELATED WORKS

According to previous researches [1-3], in comparison with male users, female ones show stronger expression of mentioned seven basic universal feeling moods.

These findings suggest women are more expected to show stronger facial expressions and report more intense mood experience than men [4].

Based on the researches related to investigation of facial expression intensity in men and women, using both ECG (signal) and questionnaire survey based approach with individuals from both genders and from different cultural background, (for seven untrained universal emotions,) there are some psychological evidences to prove that, men can censor their emotions and the facial expression has been

observed to have lower intensity for men. However, still, we need multi-racial research for gender differences, which would give us better understanding about the simultaneous effect of cultural background and gender difference on facial expression intensity.

Brody [5] in his work which was done using individuals from different countries and cultural backgrounds concluded that Women reported more powerful experiential happiness and more intense expressions related to joyful mood than did men. According to Campbell [6] female users are more likely to feel afraid than male ones and report fearful mood expression in more intense way than men.

In another study by Kirschbaum and Hellhammer [7] same statement has been reported.

In the study by Houstis and Kiliaridis [8], two-way multivariate analysis of variance was used for analyzing the facial expression's intensity of children and adults to investigate its dependence on age and gender.

According to facial expressions analysis results, they have reported that the gender differences in adults as well as age difference between groups (20 girls, 20 boys, 20 females, 20 males) are significant.

In manuscript written by Sforza, et al [9], the sex related characteristics in facial movements have been studied using database contains recorded motion details of 21 soft tissue facial landmarks, from 20 men and 20 women (from 20 to 50 years). They have concluded that, gender has statistically significant effect on total facial motion.

In the research by McDuff et al. [10], gender differences in facial mood expression, has been studied. They asked two thousands users to watch advertisements clip in home and record their own face video which contains user face with different facial responses. In their conclusion, it has been mentioned that "generally women express actions more frequently than men and in particular express more positive valence actions".

From classification point of view, gender difference was the main question of many previous researches about facial expression classification and recognition [11-14]. Several experiments were conducted on PICS Expression database [15].

Therefore, solid academic evidence suggests that users face is measurably different across the various denominations of mood and genders but the user gender's influence on the facial expressions effect on reliability of biometric system has not been investigated yet.

### III. METHODS AND MATERIALS

#### A. Database (Fig.1-a,b)

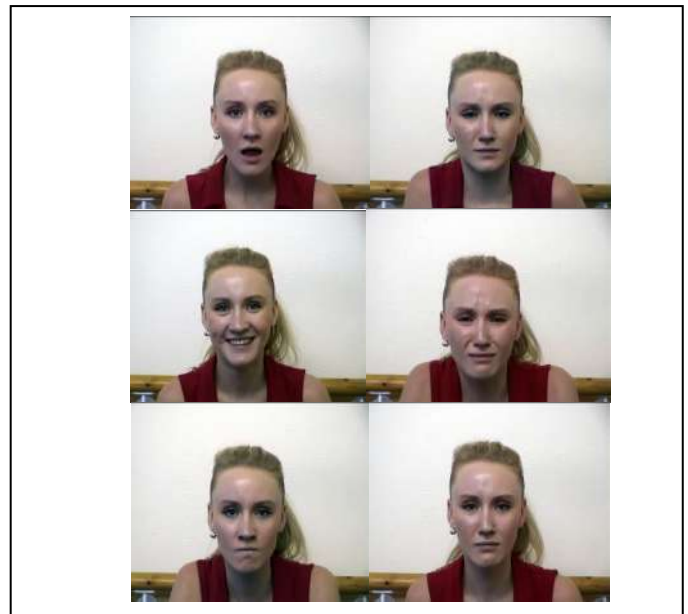
In order to find answer for this question: if the effect of facial expression on face recognition system performance evaluation results is gender dependent or not (or if for different genders, the difference between trend of comparison scores due to the facial expression would be statistically significant?), we need a proper database which must contain the face images in different moods from both female and male users. To find such database we have used Psychological Image Collection at

Stirling (PICS) database, which contains pictures from 23 users (10 male and 13 female users) in 7 original moods (2 pictures for normal, etc.) plus painful condition (10 photos in this class).

#### B. Numerical Experiment

The comparison scores (similarity score for VeriLook, and dissimilarity score for face recognition dlib), between images with neutral mood, -which is the most common and frequent mood in real life, - and images with all other moods have been obtained. So normal faces and expressive faces have been compared with each other one by one. To understand if the gender has effect on the trend of genuine scores, then we have partitionized the database in to two different sub DBs: pictures from Female users and Male users.

Fig. 1. a- Images with different moods from PICS (female



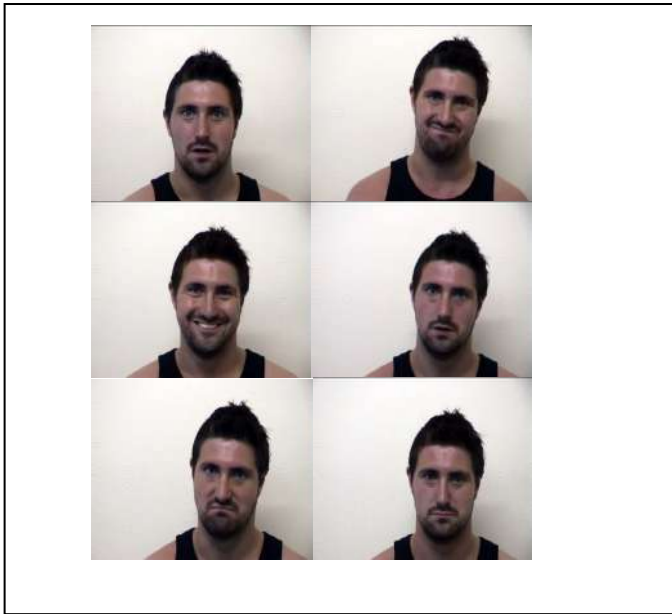
user).

Finally the comparison with mentioned strategy has been done for both of the groups (male and female users). To obtain the results two methods have been used:

1. Python Dlib for face recognition [16]: Convolutional neural network based method- the face recognition dlib gives the face distance (Euclidean distance between two 128 dimensional vector spaces) as a result of comparison. The face distance between pictures represents the dissimilarity between samples and, it is the independent variable in the interval [0, 1]. For comparison of two exactly identical pictures, the face distance will be zero, hence lower face distance means higher similarity between pictures.
2. VeriLook neurotechnology Software [17]: VeriLook is commercial matcher with an unpublished coding methodology. The matching score corresponds to the similarity of images; as a result the higher comparison score denotes a better match. Hence two cumulative distribution figures can be presented for the mentioned question.

Both of the matchers may return asymmetrical genuine matching scores and the score between the image A and the image B will remain exactly equal for the matching score between B & A; i.e. Happy vs. Anger has no difference with Anger vs. Happy).

Fig. 1. b -Images with different moods from PICS (male user).



### C. Statistical Analysis

For statistical analysis, To judge whether the observed differences in comparison scores across partitions can be considered as samples drawn from the same distribution, a two-sample Kolmogorov-Smirnov test will be applied with the significance level  $\alpha = 0.05$ . The null hypothesis  $H_0$  in test states that the samples originating from two compared partitions are drawn from the same distribution. Alternative hypothesis is:

- For python face recognition dlib:
  - $H_1: F_{\text{male}}(N\text{-all}) > F_{\text{female}}(N\text{-all})$
- For VeriLook:
  - $H_1: F_{\text{male}}(N\text{-all}) < F_{\text{female}}(N\text{-all})$

## IV. RESULTS AND DISCUSSIONS

In this section, we will discuss about the obtained results. In figure.2, the cumulative distribution function of comparison results for both groups: females (red line), males (blue one) can be illustrated. It is worth mentioning that , for presenting figure.2 we have achieved the results using first methodology called python face recognition dlib, and in figure.3, the results have been obtained by using VeriLook sdk.

As it can be illustrated in figure.2, the face distance between female normal pictures and moody pictures showed to be more than same values for men. It means that, from system perspective, due to the more facial muscle movement in women, the face shape of female users will change more than

men and as a consequence it is more probable to reject female genuine with facial expression than male users. In the other words, female face will experience more change –even deformation for more complex facial expressions- than men from biometric system point of view.

According to figure.3, both of the methods validate that the results of normal vs moody faces comparison for male users are better than female users.

Fig. 2. CDF for Male (blue), Female (Red) using python face recognition dlib.

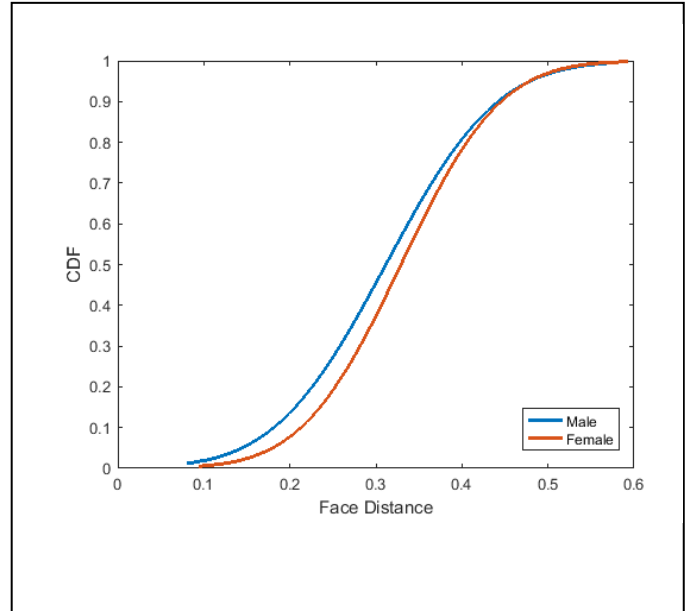
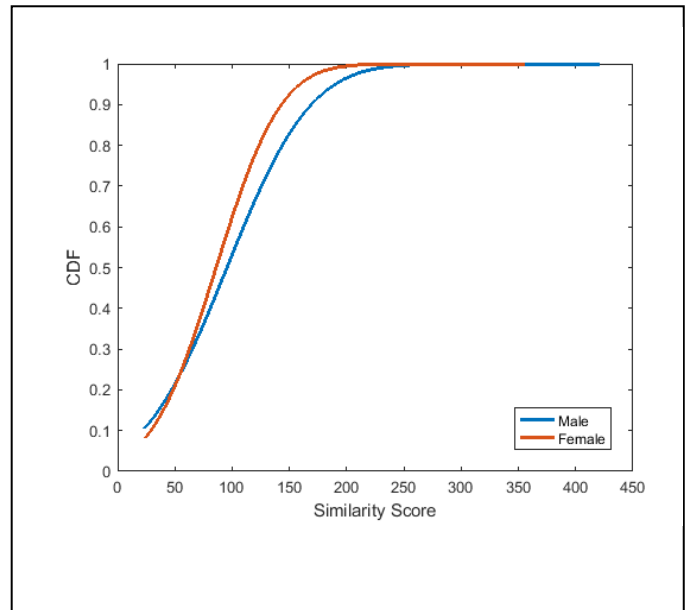


Fig. 3. CDF for Male (blue), Female (Red) using python face recognition dlib.



In table.1, the results of hypothesis tests (both Kolmogorov Simonov test and two sample t-test for both samples achieved by VeriLook and python dlib) are tabulated. The p-value in all four cases (two tests for each method) is less than 0.05 and the null hypothesis will be definitely rejected.

According to table.1 male users can be recognized easier under influence of different emotional moods, in comparison with women as psychological and classification based previous works concluded. Female users are more sensitive to emotions and as a consequence the facial muscle movement would be higher for them. Hence, it can be concluded that from biometric science point of view, the system reliability change under influence of mood must be considered gender dependent.

Table.1. results of hypothesis testing

Test→ Method↓	T-Test	KS-Test
Python dlib	0.0030	5.3373e-07
VeriLook	0.0027	0.0044

## V. CONCLUSIONS

The facial changes can affect the average genuine scores trend and consequently the reliability of the system for both genders and due to the hypothesis tests the difference is statistically significant.

To answer this question: For which gender the system under influence of emotional expressions would be more reliable, in this paper, we made an numerical experiment using PICS database and two state of the art methods called python dlib and Verilook to study the gender dependency of facial expression intense.

Final remark is:

Yes, facial expressions effects on reliability of system are different for different genders and female users show more intense expressions in comparison with men.

## ACKNOWLEDGMENT

This work is done with funding source from AMBER with sponsorship from the Marie Skłodowska-Curie EU Framework for Research and Innovation Horizon 2020, under Grant Agreement No. 675087.

## REFERENCES

- [1] H. L. Wagner, R. Buck, M. Winterbotham, "Communication of specific emotions: Gender differences in sending accuracy and communication measures," *J. Nonverbal. Behav.*, vol.17, pp.29-52, 1993.
- [2] J. G. Allen, D. M. Haccoun, "Sex differences in emotionality: A multidimensional approach," *Hum. Relat.*, vol. 29, pp. 711-722, 1976.
- [3] C. L. Barr, R. E. Kleck, Self-other perception of the intensity of facial expressions of emotion: Do we know what we show," *J. Pers. Soc. Psychol.* vol. 68, pp.608-618, 1995.
- [4] A. M. Kring, A. H. Gordon, "Sex differences in emotion: Expression, experience, and physiology," *J. Pers. Soc. Psychol.* vol. 74, pp. 686-703, 1998.
- [5] L. R. Brody, Gender, emotional expression, and parent-child boundaries. In R. D. Kavanaugh, B. Zimmerberg, & S. Fein (Eds.), *Emotion: Interdisciplinary perspectives*, 1996, pp. 139-170. Mahwah, NJ: Lawrence Erlbaum.
- [6] A. Campbell, "Staying alive: Evolution, culture and intra-female aggression," *Behav. Brain. Sci.*, vol. 22, pp. 203-252, 1999.
- [7] C. Kirschbaum, D. H. Hellhammer, "Salivary cortisol in psychoneuroendocrine research: Recent developments and applications." *psychoneuroendocrino*, vol.19, pp. 313-333, 1994.
- [8] O. Houstis, S. Kiliaridis, "Gender and age differences in facial expressions," *Eur. J. Orthod.*, vol. 31(5), pp. 459-466, 2009.
- [9] C. Sforza, A. Mapelli, D. Galante, S. Moriconi, T. Ibba, L. Ferraro, V. Ferrario, "The effect of age and sex on facial mimicry: a three-dimensional study in healthy adults," *Int. J. Oral. Maxillofac. Surg.*, vol. 39(10), pp.990-999, 2010.
- [10] D. McDuff, E. Kodra, R. Kaliouby, M. LaFrance, "A large-scale analysis of sex differences in facial expressions," *PLoS. One.*, vol. 12(4),2107.
- [11] P. Rai, P. Khanna, "Gender Classification Techniques: A Review," *Adv. Intel. Soft. Compu.*, vol. 166, pp. 51-59, 2012.
- [12] S. A. Khan, M. Nazir, S. Akram, N. Riaz, "Gender classification using image processing techniques: A survey," In *Proceedings of the 2011 IEEE 14th International Multitopic Conference (INMIC)* , pp. 25-30, 2011.
- [13] W. S. Chu, C. R. Huang, C. S. Chen, "Gender classification from unaligned facial images using support subspaces," *Inform. Sciences.*, vol. 221, pp. 98-109 , 2013.
- [14] B. Patel, R. P. Maheshwari, R. Balasubramanian, "Multi-quantized local binary patterns for facial gender classification," *Comput. Electr. Eng.*, vol. 54, pp. 271-284, 2016.
- [15] [http://pics.stir.ac.uk/2D\\_face\\_sets.htm](http://pics.stir.ac.uk/2D_face_sets.htm)
- [16] Davis E. King. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* 10, pp. 1755-1758, 2009.
- [17] [www.Neurotechnology.com](http://www.Neurotechnology.com)