Discerning Familiar/Unfamiliar Faces Using EEG

Xingyuan Guo and Ismail Jouny
Electrical and Computer Engineering, Lafayette College
Easton, Pennsylvania

***Abstract*—This paper investigates whether there is a shared pattern of EEG responses across different individuals when they look at familiar and unfamiliar faces and whether someone can deny familiarity with a certain face; if so, classification of known and unknown faces can be done across various individuals. It is shown experimentally that it's possible to detect, using EEG data, whether a subject is familiar with a face using SVM trained with pre-recorded data.**

Keywords—EEG; facial recognition; familiarity; machine learning

I. Introduction

This paper discusses how EEG data can be used as features to classify human faces based on their familiarity. It is often assumed that individuals react differently to known and unknown faces, but this study also tries to prove that there is also a shared pattern of EEG responses across individuals when they look at familiar and unfamiliar faces.

Additionally, this paper talks about data collection procedures, data processing methods and classification techniques. Though the basic principles of classification methods used including support vector machine (SVM) and extreme learning machine (ELM) will not be discussed here because they are well known in the pattern recognition community.

In related work, one study used various feature extraction methods (CSP, PCA, wavelet transform, etc.) and compared their accuracies of classifying familiar and unfamiliar faces [1], but the work mainly focused on classifying each individual's' data and comparing the effectiveness of various data processing methods. Also, another study [2] showed that using EEG data of only the first 200 msec is sufficient for the one's brain to decide whether a person looks familiar based on his/her face image. One other study [3] also showed that using combination of EEG features, classification of familiar and unfamiliar images could be done.

II. Marerials and procedures

A. *Subjects*

Two males and two females participated in this experiment.

B. *Hardware Used*

The data were collected with an Emotiv EPOC device at sampling rate of 128 Hz. The device has 14 bio-potential sensors that can wirelessly transfer raw EEG data to Matlab through a Simulink model provided by Emotiv.

C. *EEG Data Recordings*

Each of the 4 participants was shown 40 equally sized pictures of human faces – 15 of well-known familiar faces, 25 of unfamiliar faces. Each picture was displayed on a screen for 3 seconds. In total EEG data of 160 pictures were recorded. The pictures were chosen to be "familiar" or "unfamiliar" to the participants after experiments: if one "familiar" picture happened to be unfamiliar to one participant, the label (for SVM/ELM) of that picture was changed for that specific participant, and vice versa. For each of the pictures EEG data from all the 14 channels were recorded and saved for further processing.

D. *EEG Data Processing*

All 14 channels of raw EEG data of all the pictures were trimmed to maintain only the data from the first second since using only the first 200ms is sufficient for someone to tell whether someone is known or unknown based his/her face [2], then the data were filtered by a 10th order Butterworth low pass filter with cutoff frequency of 10 Hz. Then independent component analysis (ICA) was applied because ICA is considered to be capable of removing background noise [4]. The ICA part was done by using the FastICA Matlab package written by Aalto University based on a paper by Aapo Hyvärinen [5].

E. *Classification Methods*

Support vector machine (SVM) and extreme learning machine (ELM) were used to classify EEG data. SVM was commonly used for machine learning for years and ELM was a relatively new learning algorithm which was claimed to have extremely fast learning speed [6]. First, each individuals' data were separated into training and testing sets and then used to do classification. Then, data from all 160 pictures were randomly put into train/test sets to determine whether different people react to familiar and unfamiliar faces similarly for 1000 times. The results are shown in Table I and Table II. The SVM part was done by using the LIBSVM toolbox in Matlab [7]. Firstly, cross-validation was used to find the best parameters C and γ. Then the parameters were used for RBF kernel to train the SVM. For ELM, a toolbox by NTU was used, the number of hidden nodes was 20 and the activation function used was sigmoid [8]. All the data processing was done in Matlab. The data preparation, processing and classification steps are shown in Fig 1.

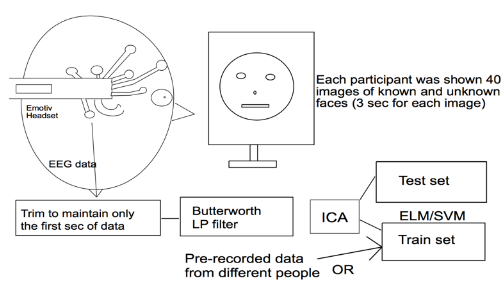

Fig. 1. Schematic diagram showing EEG data collecting, processing and classifying steps.

TABLE I. CLASSIFICATION PERFORMANCE FOR SVM AND ELM OVER EACH OF THE INDIVIDUALS' DATA

Classification methods	Subject 1 (Accuracy)	Subject 2 (Accuracy)	Subject 3 (Accuracy)	Subject 4 (Accuracy)
SVM	80%	90%	80%	76.67%
ELM	81.1%	94.28%	88.44%	80%

TABLE II. CLASSIFICATION PERFORMANCE FOR SVM AND ELM OVER THE WHOLE DATA SET (160 PICTURES)

Classification methods	Accuracy (averaged over 1000 times)
SVM	73.6238%
ELM	88.0325%

III. ANALYSIS AND CONCLUSIONS

The objective of the paper is to show a shared pattern of EEG responses across individuals when they see familiar and unfamiliar faces. As revealed in results listed above, using all 160 randomly shuffled data sets to train and test, the classification accuracy of SVM is 73.6238% and ELM with a higher accuracy of 88.0325%. It suggests that EEG data acquired from different people can be used to accurately classify familiar and unfamiliar faces so the shared pattern can be proven. Also, Table I validates that each of the individuals view familiar and unfamiliar faces differently, as suggested in previous studies [1].

For further work, it is shown in Fig. 2 that after applying digital filters and conducting ICA, the skewness of channels of EEG data display different pattern for unknown faces and known faces – skewness of unknown faces after being averaged have positive values, and of known faces after being averaged are negative. This statistically asymmetry may be explored in later studies. Also, Fig. 3 showing skewness of raw EEG data over all channels suggests that some channels display greater differences for familiar and unfamiliar faces, such as channel 2, 5 and 10.

Lastly, there are issues in this study need to be addressed: the EEG device used is only able to provide sampling rate of 128 Hz, so the quality of signal might not be satisfying enough for SVM and ELM to perform classification; moreover, the data processing methods done in the study cannot completely eliminate artifacts, such as eyes and body movements, so more noise cancelling methods need to be considered in further study to achieve better quality EEG data.

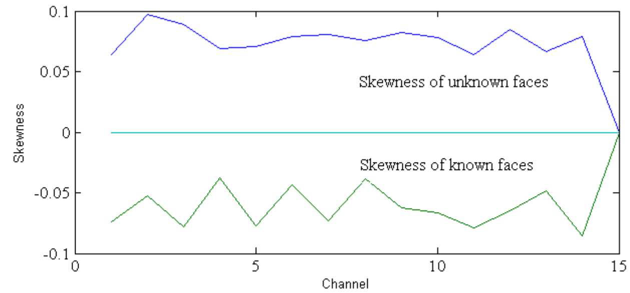

Fig. 2. Skewness of processed EEG data of familiar/unfamiliar faces averaged over all data (160 sets) over 14 channels

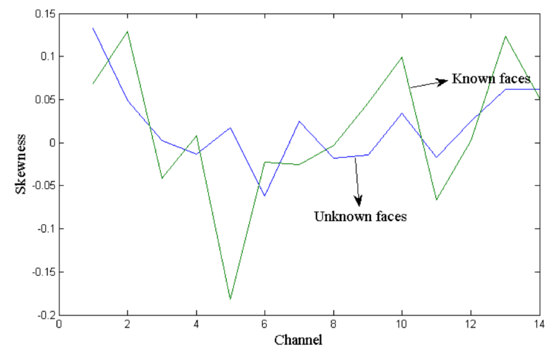

Fig. 3. Skewness of raw EEG data of familiar/unfamiliar faces averaged over all data (160 sets) over 14 channels

REFERENCES

[1] Y. Li, S. Ma, Z. Hu, J. Chen, G. Su and W. Dou, "Single trial EEG classification applied to a face recognition experiment using different feature extraction methods," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, 2015, pp. 7246-7249. doi: 10.1109/EMBC.2015.7320064S.

[2] S. Caharel, M. Ramon and B. Rossion, "Face Familiarity Decisions Take 200 msec in the Human Brain: Electrophysiological Evidence from a Go/No-go Speeded Task," in *Journal of Cognitive Neuroscience*, vol. 26, no. 1, pp. 81-95, Jan. 2014. doi: 10.1162/jocn_a_00451

[3] Z. H. E. Tan, K. G. Smitha and A. P. Vinod, "Detection of Familiar and Unfamiliar Images Using EEG-Based Brain-Computer Interface," *2015 IEEE International Conference on Systems, Man, and Cybernetics, Kowloon*, 2015, pp. 3152-3157. doi: 10.1109/SMC.2015.547

[4] Lisha Sun, Ying Liu and P. J. Beadle, "Independent component analysis of EEG signals," *Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology, 2005.*, 2005, pp. 219-222. doi: 10.1109/IWVDVT.2005.1504590

[5] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," in *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626-634, May 1999. doi: 10.1109/72.761722

[6] Guang-Bin Huang, Qin-Yu Zhu and Chee-Kheong Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, 2004, pp. 985-990 vol.2. doi: 10.1109/IJCNN.2004.1380068

[7] Chang, Chih-Chung, and Chih-Jen Lin. "LIBSVM: a library for support vector machines." *ACM Transactions on Intelligent Systems and Technology (TIST) 2*, no. 3 (2011): 27.

[8] "Basic ELM Algorithm", *Nanyang Technological University*, 2013, http://www.ntu.edu.sg/home/egbhuang/elm_codes.html