Discovering Gender Differences in Facial Emotion Recognition via Implicit Behavioral Cues

Maneesh Bilalpur*, Seyed Mostafa Kia†‡, Tat-Seng Chua§, Ramanathan Subramanian¶∥

*International Institute of Information Technology, Hyderabad, Telangana, India
†Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
‡Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, The Netherlands
§School of Computing, National University of Singapore, Singapore
¶Advanced Digital Sciences Center, University of Illinois at Urbana-Champaign, Singapore
∥School of Computing Science, University of Glasgow Singapore, Singapore

Abstract—We examine the utility of implicit behavioral cues in the form of EEG brain signals and eye movements for gender recognition (GR) and emotion recognition (ER). Specifically, the examined cues are acquired via low-cost, off-the-shelf sensors. We asked 28 viewers (14 female) to recognize emotions from unoccluded (no mask) as well as partially occluded (eye and mouth masked) emotive faces. Obtained experimental results reveal that (a) reliable GR and ER is achievable with EEG and eye features, (b) differential cognitive processing especially for negative emotions is observed for males and females and (c) some of these cognitive differences manifest under partial face occlusion, as typified by the eye and mouth mask conditions.

1. Introduction

The need to account for gender differences in interaction design and computing has spurred the evolution of the Gender HCI field [1]. Being able to recognize user demographics such as gender can benefit interactive and gaming systems in terms of (a) visual and interface design [2], (b) recommending the right games and products (via ads), and (c) providing the right motivation and feedback for enhancing user satisfaction and experience [3]. Most gender recognition (GR) systems are face or voice-based, which pose privacy concerns as face or voice samples are biometrics which enable discovery of a person’s identity.

In contrast, this work examines GR from implicit user behavioral signals such as eye movements and EEG responses. Implicit behavioral signals are invisible to the outside world and cannot be recorded without the user’s cooperation making them privacy compliant. Specifically, we attempt GR and emotion recognition (ER) captured via low-cost, off-the-shelf devices such as the Emotiv EEG and the Eyetribe eye tracker. While being easy-to-use, these devices have poor signal quality which makes analysis challenging.

Specifically, we examine gender differences by posing a facial emotion recognition task. To this end, we presented 28 viewers (14 female) with facial emotions from the Radboud face database, and asked them to recognize the presented emotion. In addition to presenting viewers with unoccluded emotive faces (no mask), we also presented them images where facial features were partially occluded, i.e., either the eyes were completely covered (eye mask), or the mouth and lower nasal region were covered (mouth mask) (see Fig.1).

The study yielded interesting results. Firstly, examination of explicit viewer responses showed that women were more accurate at recognizing especially negative emotions. Then, investigation of EEG event related potentials (ERPs) revealed differences in ERP components for positive and negative emotions in females. Analysis of eye movements showed that females were prone to fixating on the eyes for discovering emotional cues. Finally, extensive ER and GR experiments showed that (a) reliable GR (peak AUC of 0.71) and ER (peak AUC of 0.64 for recognizing positive vs negative emotions) is achievable with the considered modalities and sensors; (b) differential cognitive processing for males and females is observable while processing negative facial emotions, and (c) some of these differences manifest under occlusion as typified by the eye mask and mouth mask conditions. The next section positions our work with respect to the literature in order to highlight the motivation for this study and its novelty.

2. Related work

Many works have focused on ER with implicit behavioral signals [4, 5, 6, 7, 8], but very few works have examined on GR with such signals [9]. Also, some works have attempted to capture emotion and gender-specific differences by examining eye movement and EEG responses to emotional faces [10, 11, 12, 13]. However, none of these studies have attempted to isolate behavioral differences as well as exploit them in a computational setting. A closely related work [14] exploits gender sensitivity to mild and intense facial emotions for GR. We explicitly show gender-specific EEG and eye movement patterns characteristic of emotional face processing under occluded and unoccluded conditions, and demonstrate recognition results consistent with these findings. Spatio-temporal analysis of EEG data also reveals cognitive differences while processing facial emotions in the no mask and mask conditions.

In an exhaustive survey of GR methods [9], the authors attempt to address the GR problem via metrics like universality, distinctiveness, permanence and collectability. While bio-signals like EEG can be accurate and reliable, acquiring them is nevertheless invasive requiring express
cooperation from users. The use of low-cost, wearable devices makes our experimental design minimally invasive, our study ecologically valid and our framework suitable for user profiling at scale.

Some works have expressly attempted to recognize emotions using machine learning approaches from eye/EEG data or a combination of both. Zheng et al. [15] integrate deep belief networks with a hidden Markov model for distinguishing positive and negative emotions. Unsupervised ER from raw EEG is proposed in [16]. Zheng et al. [4] propose ER using differential entropy features based on EEG and pupil diameter. Liu et al. [5] perform ER with differential autoencoders, and attempt cross-modal ER based on shared EEG and eye-based representations. Tavakoli et al. [17] perform valence (+ve vs -ve emotion) recognition using eye movements with an emphasis on evaluating eye-based features and their fusion, and achieve 52.5% accuracy with discriminative feature selection and a linear SVM.

Upon examining the literature, we summarize the research contributions of this work as follows: (1) It is one of the very few works to computationally perform GR and ER based on EEG and eye-based features; we also supplement the recognition results with explicit user behavioral data and isolation of EEG and eye-based patterns; (2) We use low-cost, off-the-shelf EEG and eye movement sensors to this end as against lab-grade equipment used by most works, which affirm the ecological validity of our findings and the feasibility of employing the proposed framework for large-scale user profiling; (3) Use of the mask vs no mask paradigms allows for a fine-grained examination of gender differences. Some gender-specific behaviors (e.g., proclivity to fixate on the eyes for women or gender information encoded in the EEG data at specific electrodes) manifest under the eye and mouth mask conditions. The following section describes our experimental set up and protocol.

3. Materials and Methods

We examined gender differences in visual emotional face processing when a) the faces were fully visible, and (b) some facial features were occluded. Specifically, we investigated gender sensitivity to emotions when the eye and mouth regions are occluded (see Fig.1), as the importance of these features towards conveying facial emotions has been highlighted by prior works [10, 13]. The three conditions are denoted as the no mask, eye mask and mouth mask conditions throughout the paper.

Participants: 28 subjects of different nationalities (14 male, age 26.1 ± 7.3 and 14 female, age 25.5 ± 6), with normal or corrected vision, took part in our study.

Stimuli: We used emotional faces of 24 models (12 male, 12 female) from the Radboud Faces Database (RaFD) [19]. RaFD includes facial emotions of 49 models rated for clarity, genuineness and intensity, and the 24 models were chosen such that their Ekman facial emotion ratings were roughly matched based on these ratings. We then morphed the emotive faces from neutral (0% intensity) to maximum (100% intensity) to generate intermediate morphs in steps of 5%. Emotional morphs with 55–100% intensity were used in this work. Eye and mouth-masked faces were generated upon automatically locating facial landmarks via [20]. The eye mask covered the eyes and nasion, while the mouth mask covered the mouth and the nose ridge. The (un)masked faces were 361 × 451 pixels in size, encompassing a visual angle of 9.1° and 11.4° about x and y at 60 cm screen distance.

Protocol: The protocol is outlined in Fig.1 and involved the presentation of unmasked and masked faces to viewers over two separate sessions, with a break in-between to avoid fatigue. We chose one face per model and emotion, resulting in 144 face images (1 morph/emotion × 6 emotions × 24 models). In the first session, these faces were shown as is in random order (no mask condition), and were re-presented with an eye or mouth mask (eye/mouth mask conditions) in the second session. We ensured a 50% split of the eye/mouth-masked faces, and the masked faces were also shown randomly to viewers.

During each trial, the (un)masked face was displayed for 4s preceded by a fixation cross for 500 ms. The viewer then had a maximum of 30s to make one out of seven choices concerning the facial emotion (6 emotions plus neutral) via a radio button. Neutral faces were only utilized for morphing purposes and not used in the experiment. During the experiment, viewers’ EEG signals were acquired via the 14-channel Emotiv epoc device, and eye movements were recorded with the Eyetribe tracker. The experiment was split into 4 segments to minimize acquisition errors, and took about 90 minutes.

3.1. User Responses

We first compare male and female sensitivity to emotions based on response times (RTs) and recognition rates (RRs), and then proceed to examine their implicit eye movements and EEG responses. RT and RR denote performance measures computed from explicit viewer responses.

Response Time (RTs): Overall RTs for the no mask, eye mask and mouth mask conditions were respectively found to be 1.44 ± 0.24, 1.17 ± 0.12 and 1.25 ± 0.09 seconds, implying that facial ER was fairly instantaneous, and that

1. Anger, Disgust, Fear, Happy, Sad and Surprise.
viewer responses were surprisingly faster with masked faces. A fine-grained comparison of male and female RTs in the three conditions (Fig.2(a)) revealed that females ($\mu_{RT} = 1.40 \pm 0.10s$) were generally faster than males ($\mu_{RT} = 1.60 \pm 0.10s$) at recognizing unmasked emotions. Female alacrity nevertheless decreased for masked faces, with males responding marginally faster for eye masked faces ($\mu_{RT}(\text{male}) = 1.13 \pm 0.11s$ vs $\mu_{RT}(\text{female}) = 1.21 \pm 0.13$s), and both genders responding with similar speed for mouth masked faces ($\mu_{RT}(\text{male}) = 1.24\pm0.10s$ vs $\mu_{RT}(\text{female}) = 1.25\pm0.09$s).

**Recognition Rates (RRs):** Overall, RRs for unmasked emotions ($\mu_{RR} = 77.6$) were expectedly higher than for eye-masked ($\mu_{RR} = 59.7$) and mouth-masked ($\mu_{RR} = 63.5$) emotions. Happy faces were recognized most accurately in all three conditions. Specifically focusing on gender differences (Fig.2(b)), females recognized facial emotions more accurately than males and this was particularly true for negative (A,D,F,S) emotions— male vs female RRs for these emotions differed significantly in the no mask ($\mu_{RR}(\text{male}) = 54.3$ vs $\mu_{RR}(\text{female}) = 61.2$, $p<0.05$) and eye mask conditions ($\mu_{RR}(\text{male}) = 51.8$ vs $\mu_{RR}(\text{female}) = 58.1$, $p<0.05$), and marginally for the mouth mask condition ($\mu_{RR}(\text{male}) = 52$ vs $\mu_{RR}(\text{female}) = 55.8$, $p = 0.08$).

4. Analyzing Implicit Behavioral Cues

We first describe the methodology employed to extract brain and eye-related features, before discussing emotion and gender recognition with the extracted features.

4.1. EEG preprocessing

We extracted epochs for each trial (4.5s of stimulus-plus-fixation viewing time @ 128 Hz sampling rate), and the 64 leading pre-stimulus samples were used to remove DC offset. This was followed by (a) EEG band-limiting to within 0.1–45 Hz, (b) Removal of noisy epochs via visual inspection, and (c) Independent component analysis (ICA) to remove artifacts relating to eye-blinks, and eye and muscle movements. Muscle movement artifacts in EEG are mainly concentrated in the 40–100 Hz band, and are removed upon band-limiting and via inspection of ICA components. Finally, a 7168 dimensional (14×4×128) EEG feature vector was generated from concatenation of the 14 EEG channels over 4s of stimulus viewing.

4.1.1. ERP analysis. Event Related Potentials (ERPs) are (averaged) time-locked neural responses related to sensory and cognitive events. ERPs occurring in the first 100ms post stimulus presentation are stimulus-related (or exogenous), while later ERPs are cognition-related (or endogenous). We examined the first second of EEG epochs for ERPs characteristic of emotion and gender.

In a related work studying gender differences in emotional processing, Lithari et al. [21] observed enhanced N100 and P300 ERP components in females for negative stimuli. Fig.3 shows female ERP signals observed in the occipital O2 electrode for the no mask and eye mask conditions. The occipital lobe represents the visual processing center in the brain. Clearly, one can note enhanced N100 and P300 ERP components for negative emotions in the no mask condition (Fig.3(a)). This effect is attenuated in the eye mask (Fig.3(b)) and mouth mask conditions, although one can note stronger N400 amplitudes for anger and disgust in the eye mask case. Such a pattern was not observable from male ERPs. Overall, the observed ERPs affirm that gender differences in emotional processing can be captured with the low-cost Emotiv EEG device.

4.2. Gaze data preprocessing

Raw gaze data acquired from the EyeTribe device at 30 Hz sampling rate were used to extract fixations with 100 ms temporal threshold via the EyeTribe toolbox. Saccades were assumed as the transitions between fixations. We also computed the duration fixated by male and female viewers for the six facial regions, namely, nose, eyes, mouth, forehead, cheeks and chin. The normalized fixation duration distribution over these six regions for the three conditions is presented in Fig.4(a–c).

Fig.4(a–c) show that the eye, mouth and nasal face regions attract maximum visual attention during facial emotion recognition, in line with prior studies such as [10, 18]. Prior works such as [22] have also observed that females look up to the eyes for emotional cues, which is mirrored by longer female fixations around the eyes in the no mask and mouth mask conditions. Fig.4(b) suggests that when eye information is unavailable, females tend to concentrate on the mouth and nasal regions for emotional cues. Upon extracting the fixations and saccades, we adopted features used for recognizing valence (pleasant vs unpleasant emotions) in [17], namely, saccade orientation, top–ten salient coordinates, saliency map and histograms of a) saccade slope, b) saccade length, c) saccade velocity, d) first saccade duration, and e) last saccade duration.
4.3. Experiments and Results

This section presents emotion recognition (ER) and gender recognition (GR) results obtained with EEG features, eye-based features and their combination, under the three considered conditions. Recognition experiments were performed on data compiled from trials where viewers correctly recognized the presented facial emotion.

4.3.1. Experimental settings. We considered (i) EEG features, (ii) attention-based features, (iii) concatenation of the two (termed early fusion or EF), and (iv) probabilistic fusion of the EEG and eye-based classifiers (late fusion or LF) for our analyses, in the three conditions during our recognition experiments. For LF, we employed the $W_{ext}$ technique [23] briefly described as follows. From the EEG and eye-based classifier outputs, the test set posterior probability for class $j$ is computed as $\sum_{i=1}^{2}\alpha_i^{*}t_i^{j}p_i^{j}$, where $i$ indexes the two modalities, $p_i^{j}$'s denote posterior classifier probability for class $j$ with modality $i$ and $\{\alpha_i^{*}\}$ are the optimal weights maximizing test performance, as determined via a 2D grid search. If $F_i$ denotes the training performance for the $i^{th}$ modality, then the normalized training performance $t_i = \alpha_i F_i / \sum_{i=1}^{2}\alpha_i F_i$ for given $\alpha_i$.

We considered four classifiers including Naïve Bayes (NB), Linear Discriminant Analysis (LDA), Support Vector Machines with RBF (RSVM) and linear (LSVM) kernels in our recognition experiments. We adopted the area under ROC curve (AUC) metric for performance evaluation. AUC measures area under the ROC curve, which plots the true positive rate against false positive rate with varying threshold. A random classifier would generate an AUC score of 0.5, while a perfect classifier would return an AUC of 1. Also, as we attempt recognition with few training data, we report ER/GR results over five repetitions of 10-fold cross validation (CV). CV is typically used to overcome the overfitting problem on small datasets, and optimal SVM parameters were determined from the range $[10^{-4}, 10^4]$ via an inner 10-fold CV on the training set.

4.3.2. Valence (+ve vs -ve emotion) Recognition. As explicit behavioral responses (Sec.4.1.[7]) and ERP-based (Sec.4.1.[7]) analyses suggest that (a) females are especially sensitive to negative facial emotions, and (b) there are cognitive differences in the processing of positive and negative emotions for females, we attempted to classify positive (H, Su) and negative (A,D,F,S) emotions. Fig.4[d] presents ER results obtained with the LDA classifier for male and female data, with unimodal and multimodal features for the three conditions. The best ER results are expectedly achieved for the no mask condition, and with eye-based features (AUC = 0.63 for females, and 0.64 for males), while minimal recognition is noted for the eye mask condition reiterating the importance of eyes towards encoding emotional cues. EEG-based features only produce slightly better than chance recognition with a peak AUC score of 0.54. Multimodal recognition is slightly better than unimodal, and higher AUC scores are achieved in the masked conditions suggesting the complementarity of EEG and eye-based features when only partial facial information is available.

While observed ER results are modest, they are still better or competitive with respect to prior eye and neural based ER approaches. Eye based features are found to achieve 52.5% valence recognition accuracy in [17], where emotions are induced in viewers by presenting a diverse types of emotional scenes to viewers as against emotional faces specifically employed in this work. Also, prior neural-based emotional studies [6], [24] achieve only around 60% valence recognition with lab-grade sensors. On the contrary, implicit behavioral cues are acquired with low-cost EEG and eye tracking devices in this work.

4.3.3. Gender Recognition. Very few works have investigated GR via EEG and eye-based responses. Table 1 presents GR results with features extracted for (a) all faces, and (b) emotion-specific faces for the three conditions. The classifier employed for GR is also specified with parentheses. As with ER, the best GR performance (AUC = 0.71) is observed for the no mask condition but with EEG features. With both EEG and eye-based cues, slightly higher GR performance is noted for the eye mask condition as compared to the mouth mask condition considering viewer responses for all as well as emotion specific faces. Overall, EEG features appear to encode gender differences better than eye-based features. Also, early or late fusion of the two modalities does not improve GR performance by much, even though one can note that fusion performs better than at least one of the modalities in the eye mask and mouth mask conditions.

A close examination of the emotion-wise results reveals that optional GR is mainly achieved with viewer responses observed for negative facial emotions. This again is consistent with the observations from explicit user responses (Sec.4.1.), where females were found to recognize negative facial emotions better than males. Overall, the obtained results point to systematic differences between males and females in the cognitive processing...
of emotional faces. Focusing on the standard deviation (sd) in AUC scores for emotion-wise GR, one can note much higher sd values in the eye mask and mouth mask conditions as compared to no mask, indicating the difficulty in recognizing facial emotions under occlusion conditions. However, relatively lower sd values are noted with LF implying that probabilistic fusion enables better exploitation of the complementarity in the eye and brain-based signals, resulting in more robust GR.

4.3.4. Spatio-temporal EEG analysis. As EEG represents multi-channel time-series data, we performed spatial and temporal analyses on the EEG signals to discover (i) the critical time window for GR over the 4s of stimulus presentation, and (ii) the key EEG electrodes encoding gender differences in emotional face processing. 

Fig 5(a–c) show the temporal variance in GR performance for three considered conditions when the 4s presentation time is split into four 1s-long windows. These results are computed considering emotion-specific EEG epochs using RSVM. No significant difference in GR performance is observed over time for the no mask scenario, and an AUC score of around 0.65 can be consistently noted over time. Barring exceptions, the GR AUC scores for the eye mask and mouth mask conditions are lower than for the no mask condition. Also, maximum GR is predominantly noted for later temporal windows for mask conditions, implying that the discriminative cognitive processing between males and females happens later in time when only partial emotional information is made available to viewers. For the eye mask scenario, maximum GR AUC scores are achieved for anger (AUC = 0.71 in W3), disgust (AUC = 0.65 in W1) and surprise (AUC = 0.65 in W2, W3 and W4). For the mouth mask condition, highest AUCs are achieved for anger (AUC = 0.76 in W1 and 0.78 in W4), disgust (AUC = 0.68 in W2) and happy (AUC=0.69 in W2). However, low GR performance is generally noted over all time windows for the fear and surprise emotions in this condition, attributable to the general difficulty in recognizing these emotions when mouth-related information is unavailable.

Results of spatial analysis to identify the (top three) electrodes that best encode gender differences in emotion processing are shown in Fig 5(d). In the no mask condition, very similar AUC scores are observed for the frontal F8 (0.55), AF3 (0.52) and T7 (0.52) electrodes.

Much higher numbers are noted for the eye mask and mouth mask conditions, suggesting that region-specific differences in cognitive processing for males and females manifest under occlusion conditions. For the eye mask condition, highest AUCs are noted for O2 (0.87), P7 (0.81) and FC6 (0.76), while for mouth mask F3 (AUC = 0.89), P7 (AUC = 0.87) and O1 (AUC = 0.87). An interesting observation is that a majority of the electrodes are located on the left hemisphere of the brain (see Emotiv electrode configuration for details).

5. Discussion and Conclusion

This work expressly examined the utility of implicit behavioral cues, in the form of EEG brain signals and eye movements to encode gender-specific information during visual processing of emotional faces. In addition to the use of unoccluded (no mask) emotive faces, we also employed occluded (eye and mouth masked) faces to this end.

The study yielded several interesting results. Firstly, examination of explicit viewer responses revealed the higher sensitivity of females to negative emotions. We then proceeded to examine their implicit behavioral signals, which revealed (a) stronger N100, P300 and N400 female ERP components for negative emotions, and (b) the proclivity of females to fixate on the eyes for discovering emotional cues.

Emotion and gender recognition experiments with EEG and eye-based features, as well as their combination revealed that optimal recognition of emotional valence is achieved with eye-based features, while best GR performance is achieved with EEG. Also, best GR with emotion-specific features is noted for negative emotions consistent with the analysis involving explicit viewer responses. EEG features generally outperform eye-based features for the purpose of GR. Also, the fusion of EEG and eye-based features is not significantly beneficial even though more robust classification results are achieved via late fusion of unimodal results, suggesting that the EEG (neural) and eye (physiological) signals encode complementary gender-specific information. Finally, spatio-temporal analysis of EEG data revealed that superior GR is achieved with EEG from later time windows in the mask conditions, and that frontal and occipital electrodes best encode gender differences in cognitive processing of emotional faces.

The obtained results are competitive in terms of ER and GR performance, especially considering that implicit user behavioral cues are acquired with low cost sensors,
Figure 5. (a–c): Temporal variance in GR performance for the no mask, eye mask and mouth mask conditions. (d) Top three electrodes showing best GR performance for the three conditions.

and also consistent with prior observations. Lithari et al. [21] noted amplitude and phase-related differences for positive and negative emotions in males and females, and some of these findings are mirrored in our ERP analyses. Several eye tracking works (e.g., see [25]) have noted enhanced female sensitivity for negative emotions, while other neural studies [26] have noted that gender differences in emotion processing are best encoded by the frontal and occipital brain lobes. As reliable GR (max AUC = 0.71) is achievable with the considered features, future work will explore use of these modalities for user profiling in gaming and interactive scenarios to facilitate applications like targeted advertising. Also, recent developments in the field of deep learning can be exploited to discover latent feature spaces that are robust to noise and best capture gender-based cognitive impressions.

Acknowledgement
This study is supported by the Human Centered Cyber-physical Systems research grant from Singapore’s Agency for Science, Technology and Research (A*STAR).

References
[1] S. Wiedenbeck, V. Grigoreanu, L. Beckwith, and M. Burnett, “Gender HCI: What about the software?” Computer, vol. 39, pp. 97–101, 2006.
[2] M. Czerwinski, D. S. Tan, and G. G. Robertson, “Women take a wider view,” in Conference on Human Factors in Computing Systems (CHI), 2002, pp. 195–202.
[3] J. D. Schwark, I. Dolgov, D. Hor, and W. Graves, “Gender and personality trait measures impact degree of affect change in a hedonic computing paradigm,” International Journal of Human Computer Interaction, vol. 29, no. 5, pp. 327–337, 2013.
[4] W.-L. Zheng, B.-N. Dong, and B.-L. Lu, “Multimodal emotion recognition using EEG and eye tracking data,” in Engineering in Medicine and Biology Society (EMBC), 2014, pp. 5040–5043.
[5] W. Liu, W.-L. Zheng, and B.-L. Lu, “Multimodal emotion recognition using multimodal deep learning,” arXiv preprint arXiv:1602.08225, 2016.
[6] M. Abadi, R. Subramanian, S. Kia, P. Avesani, I. Patras, and N. Sebe, “DECAF: MEG-based multimodal database for decoding affective physiologival responses,” IEEE Transactions on Affective Computing, vol. 6, no. 3, pp. 209–222, 2015.
[7] S. Koelstra, C. Mohl, M. Soleymani, I.-S. Lee, A. Yazdan, T. Ebrahimi, T. Pan, A. Nigholi, and I. Patras, “DEAP: A Database for Emotion Analysis Using Physiological Signals,” IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, 2012.
[8] R. Subramanian, J. Wache, M. Abadi, R. Vierru, S. Winkler, and N. Sebe, “ASCERTAIN: Emotion and personality recognition using commercial sensors,” IEEE Transactions on Affective Computing, 2016.
[9] Y. Wu, Y. Zhuang, X. Long, F. Lin, and W. Xu, “Human gender classification: A review,” arXiv preprint arXiv:1507.05122, 2015.
[10] M. W. Schurgin, J. Nelson, S. Iida, H. Ohura, J. Y. Chiao, and S. L. Franconeri, “Eye movements during emotion recognition in faces,” Journal of Vision, vol. 14, no. 13, pp. 1–16, 2014.
[11] M. D. Zotto and A. J. Pegna, “Processing of masked and unmasked emotional faces under different attentional conditions: an electrophysiological investigation,” Frontiers in psychology, vol. 6, p. 1691, 2015.
[12] H. Katti, R. Subramanian, M. Kankanhilli, N. Sebe, T.-S. Chua, and K. R. Ramakrishnan, “Making computers look the way we look: exploiting visual attention for image understanding,” in ACM Int’l conference on Multimedia, 2010, pp. 667–670.
[13] Subramanian, D. Shinar, N. Sebe, and D. Melcher, “Emotion modulates eye movement patterns and subsequent memory for the gist and details of movie scenes,” Journal of vision, vol. 14, no. 3, pp. 1–18, 2014.
[14] M. Bilalpur, S. M. Kia, M. Chawla, T.-S. Chua, and R. Subramanian, “Gender and emotion recognition with implicit user signals,” in International Conference on Multimodal Interaction, 2017.
[15] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, “EEG-based emotion classification using deep belief networks,” in International Conference on Multimedia and Expo (ICME), 2014, pp. 1–6.
[16] S. Juraycharenol, T. Pun-Ngum, and P. Isarasena, “EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation,” The Scientific World Journal, vol. 2014, 2014.
[17] H. R.-Tavakoli, A. Atyabi, A. Rantanen, S. J. Laukka, S. Nefti-Meziani, and J. Heikkil, “Predicting the valence of a scene from observers eye movements,” PLoS ONE, vol. 10, no. 9, pp. 1–19, 2015.
[18] R. Subramanian, V. Yanulevskaya, and N. Sebe, “Can computers learn from humans to see better?: Inferring scene semantics from viewers’ eye movements,” in ACM International Conference on Multimedia, 2011, pp. 33–42.
[19] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg, “Presentation and validation of the raudbod faces database,” Cognition and Emotion, vol. 24, no. 8, pp. 1377–1388, 2010.
[20] T. Baltrusaitis, P. Robinson, and L.-P. Morency, “Openface: an open source face analysis toolkit,” in IEEE Winter Conference on Applications of Computer Vision, 2016.
[21] C. Lithari, C. A. Frantzidis, C. Papadelis, A. B. Vivas, M. A. Klados, C. Kourtidou-Papadeli, C. Pappas, A. A. Ioannides, and P. D. Bamidis, “Are females more responsive to emotional stimuli? a neurophysiological study across arousal and valence dimensions,” Brain Topography, vol. 23, no. 1, pp. 27–40, 2010.
[22] L. J. Wells, S. M. Gillespie, and P. Rothstein, “Identification of emotional facial expressions: Effects of expression, intensity, and sex on eye gaze,” PloS one, vol. 11, no. 12, 2016.
[23] S. Koelstra and I. Patras, “Fusion of facial expressions and EEG for implicit affective tagging,” Image and Vision Computing, vol. 31, no. 2, pp. 164–174, 2013.
[24] C. Mulh, B. Allison, A. Nigholt, and G. Chanel, “A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges,” Brain-Computer Interfaces, vol. 1, no. 2, pp. 66–84, 2014.
[25] J. A. Hall and D. Matsumoto, “Gender differences in judgments of multiple emotions from facial expressions,” Emotion, vol. 4, no. 2, pp. 201–206, 2004.
[26] E. B. McClure, C. S. Monk, E. E. Nelson, E. Zarahn, E. Liebenluft, R. M. Bilder, D. S. Charney, M. Ernst, and D. S. Pine, “A developmental examination of gender differences in brain engagement during evaluation of threat,” Biological psychiatry, vol. 55, pp. 1047–55, 2004 Jun 1 2004.