Emotional Reaction Recognition from EEG

Kiret Dhindsa
School of Scientific Computing and Engineering
McMaster University
Hamilton, Ontario, Canada, L8S 4L8
Email: dhindsj@mcmaster.ca

Suzanna Becker
Department of Psychology, Neuroscience, and Behaviour
McMaster University
Hamilton, Ontario, Canada, L8S 4L8
Email: becker@mcmaster.ca

***Abstract*—In this study we explore the application of pattern recognition models for recognizing emotional reactions elicited by videos from electroencephalography (EEG). We show that both the presence and magnitude of each emotion can be predicted above chance levels with up to 88% accuracy. Furthermore, we show that there are differences in classifiability for different emotions and participants, but whether a participant's data can be classified with respect to different emotions can itself be predicted from their EEG. Index Terms– Emotion recognition, electroenecephalography (EEG), pattern recognition, classification, regression, individual differences, affective computing**

I. Introduction

The ability to recognize emotional reactions based on biological data has a wide variety of applications [1], including in therapy for mood disorders and in prosthetic devices for those with communication disorders. In particular, the rapidly growing field of affective computing is focused on integrating the emotional states of users into computer applications [2], [3]. However, more accurate pattern recognition is needed to make such applications reliable.

Machine learning is a quickly expanding field that trains data-driven pattern recognition models [4]. In recent years, machine learning approaches have been used to classify emotional imagination [5], and emotional reactions to pictures [6], audio samples [7], and videos [8] from electroencephalography (EEG). However, these studies have typically been limited to classifying only positive versus negative emotions, and use research-grade EEG hardware which is costly and impractical for most real-world applications.

In this study, we extend previous results on classification of emotional reactions and classify high versus low emotional experience with respect to 11 different emotions using a consumer-grade EEG headband. In addition, we show that we can predict from the same EEG signals whether inter-subject classification will result in above-chance classification at all. Finally, we extend our results beyond simple binary classification and employ machine learning regression analysis to predict the magnitude of the emotional reaction as well.

II. Data Acquisition

A. Participants and Materials

Forty undergraduate students from McMaster University (24 female) participated in the study. No exclusion criteria were applied. Participants provided informed consent and the study was approved by the McMaster Research Ethics Board.

A total of 102 videos acquired through Youtube were used. The videos varied in length from 51s to 300s with an average length of 138.7s. Approximately half of the video clips were from Hollywood movies, and the other half were taken in real life situations using video cameras.

EEG was recorded with the Muse headband [9]. The Muse headband has four EEG electrodes located at T9, Fp1, Fp2, and T10 according to the International 10-20 system. The reference electrode is situated at Fpz with DRLs (driven right legs) on both sides. EEG was sampled at 220 Hz with 50 Hz and 60 Hz notch filters implemented in hardware.

B. Experimental Protocol

Data were collected using Matlab R2013b [10] and stimuli were presented using the Psychophysics Toolbox [11]. Participants were set up with an EEG headband and seated in front of a laptop in a private room. The participants watched approximately 60 minutes of videos, amounting to 17.2 $\pm$ 3.5 videos watched per subject. After each video, participants rated on a scale of one to ten the extent to which they experienced the following emotions in response to the video they just watched: 'Interest', 'Amusement', 'Happiness', 'Sadness', 'Fear', 'Disgust', 'Anger', 'Hope', 'Relief', 'Surprise', and 'Sympathy'. Videos were pre-labelled as positive or negative and an approximately equal number of positive versus negative videos were played for each participant.

III. Data Analysis

A. Preprocessing

EEG corresponding to each video was de-meaned and cleaned of artifacts using the Filter-Bank Artifact Rejection (FBAR) toolbox [12]. For classification analysis, self-reported emotional ratings were transformed into binary labels by grouping ratings from 1 to 5 in one class and ratings from 6 to 10 in another class. Emotional ratings were kept in their original form for regression analysis.

B. Feature Extraction

Fourth order Butterworth bandpass filters were used to extract the theta (4-7 Hz), alpha (8-12 Hz), beta (13-30 Hz), and gamma (31-45 Hz) bands from each of the four EEG channels.

From each of these frequency bands, a wide variety of features were computed in order to increase the likelihood that a subset of features would be found to be useful for each emotion and for each task, including spectral and cross-spectral density, coherence, cross-frequency coupling, bispectrum, bicoherence, quadratic phase coupling, and alpha asymmetries. In total, 279 features were computed from the EEG recording during each video viewing.

1) Power Spectal Density: For each of the four channels, the power spectral density (PSD) was estimated for the theta, alpha, and beta bands using the definition

$$P_{XX}(f) = \lim_{T\to\infty} \mathbf{E}\left[|\hat{X}_T(f)|^2\right] \tag{1}$$

where $\hat{X}_T(\omega)$ is the finite-time Fourier transform of a signal X. This yielded 12 features.

2) Cross-Spectrum: The cross-spectral density (CSD) extends the PSD to the Fourier transform of two signals, X and Y and is defined as

$$P_{XY}(f) = \lim_{T\to\infty} \mathbf{E}\left[\hat{X}_T(f)\hat{Y}_T(f)\right]. \tag{2}$$

The CSD was computed for the theta, alpha, beta, and gamma bands for all pairs of channels, resulting in 24 features.

3) Coherence: Coherence between all pairs of electrodes within the theta, alpha, beta, and gamma bands was computed. Coherence between two signals X and Y is defined as

$$C_{XY}(f) = \frac{|P_{XY}(f)|^2}{P_{XX}(f)P_{YY}(f)}, \tag{3}$$

where P_{XX} is the PSD as defined in Eq. 1, and P_{XY} is the CSD as defined in Eq. 2. In total, 24 coherence features were computed.

4) Cross-Frequency Coupling: The weighted phase locking factor (WPLF) [13] was used as a measure of cross-frequency coupling. WPLF is a measure of coupling strength and the preferred phase angle for a pair of signals. WPLF is given by

$$WPLF = \frac{1}{T}\sum_{t=1}^{T} e^{i\theta(t)}, \tag{4}$$

where T is time and θ is the phase difference between two signals. WPLF magnitude and phase angle was computed for each pair of channels and for each pair of frequency bands, resulting in 96 features.

5) Bispectrum, Bicoherence, and Quadratic Phase Coupling: The bispectrum of a signal is the 2D Fourier Transform of the third order cumulant generating function and is given by

$$B(f_1, f_2) = \mathcal{F}^*(f_1 + f_2)\mathcal{F}(f_1)\mathcal{F}(f_2), \tag{5}$$

where $\mathcal{F}$ is the Fourier Transform and $\mathcal{F}^*$ is its complex conjugate [14].

Bicoherence is the normalized bispectrum for a signal in n bins (32 bins were used here):

$$B_c(f_1, f_2) = \frac{|\sum_n \mathcal{F}_n(f_1)\mathcal{F}_n(f_2)\mathcal{F}_n^*(f_1+f_2)|}{\sum_n |\mathcal{F}_n(f_1)\mathcal{F}_n(f_2)\mathcal{F}_n^*(f_1+f_2)|}. \tag{6}$$

Quadratic Phase Coupling features are obtained from the autoregressive parameters of the bicoherence of a signal. These parameters provide information about the change in bicoherence over the frequency landscape. A 6$^{\text{th}}$ order autoregressive analysis was performed here.

The bispectrum, bicoherence and QPC are used to capture non-linear interactions between frequency pairs in a signal with respect to magnitude, frequency-normalized magnitude, and phase [14]. The sum and sum-of-squares of the bispectrum and bicoherence for the theta, alpha, and beta bands of each channel, and their corresponding QPC autoregressive coefficients were extracted as features, for a total of 120 features.

6) Alpha Asymmetries: Asymmetric alpha power in the frontal and temporal cortices is associated with emotional responses to video [15]. Three alpha asymmetry features were computed: frontal alpha asymmetry (FAA) using only the two frontal channels, temporal alpha asymmetry (TAA) using only the two temporal channels, and global alpha asymmetry (GAA) using all four channels. Alpha asymmetries were computed using the formula

$$AA = \frac{L - R}{L + R} \tag{7}$$

where L and R are left and right alpha power respectively.

C. Feature Selection

From the 279 features extracted, small subsets were selected for classification using the minimum-Redundancy Maximum-Relevance (mRMR) feature selection method [16] to reduce the risk of overfitting. The mRMR method is an information theoretic approach to feature selection which aims to maximize the mutual information between the subset of selected features and the true training labels while simultaneously minimizing the mutual information among selected features. The number of features to select is a parameter chosen via cross-validation from the set $K \in \{4, 8, 16, 32\}$. Features were selected using training data only.

D. Classification and Regression

Two classifiers were used in this study: a support vector machine with a radial basis function kernel ($C = 1$, $\sigma_{RBF} = 1$) [17] and logistic regression (LR) with elastic net regularization ($\alpha = \frac{1}{4}$) [18]. The parameter λ was chosen with a nested validation set comprising 10% of the training data. Regression analyses were performed using a boosted decision tree with 500 nodes [19].

E. Classification and Regression Tasks

The EEG data recorded during each video were used in five different tasks. In Task 1, leave-one-subject-out cross-validation (LOSO-CV) was performed by using each participant's data as a test set once while the remaining participants for training. In Task 2, within-subject analysis was performed with leave-one-video-out cross-validation (LOVO-CV). In Task 3, LOSO-CV was performed again but this time using only the participants who had classifiable data (greater than 60% classification accuracy or statistically greater than

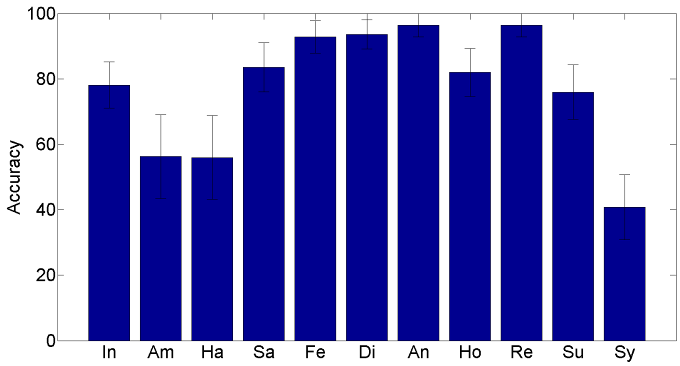

Fig. 1: Prediction rate of classifiability of participants with respect to each emotion. In = Interest, Am = Amusement, Ha = Happiness, Sa = Sadness, Fe = Fear, Di = Disgust, An = Anger, Ho = Hope, Re = Relief, Su = Surprise, Sy = Sympathy.

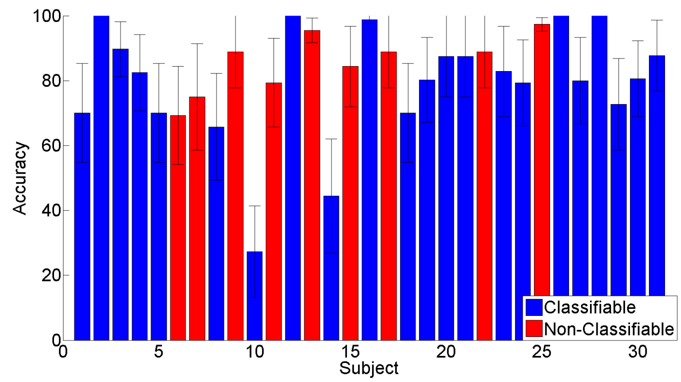

Fig. 2: Prediction rate of Classifiability of each participant averaged across all emotions.

zero correlation between predicted and observed emotional response) in Task 2. In Task 4, data from all participants were used together to perform regular 10-fold cross-validation with randomized partitioning of training and test sets with a 75%-25% split. Finally, in Task 5, 10-fold cross-validation was performed using only the participants who had classifiable data in Task 2.

IV. Results

Table I shows the classification accuracies for each task and each emotion. The dominant relevant features were alpha asymmetries, frontal alpha and beta bicoherence, temporal theta coupling, and alpha and beta fronto-temporal WPLF. Similarly, the results of regression analysis are shown in Table II. Regression results were driven almost exclusively by beta band QPC features in the frontal channels.

There was a high degree of inter-subject variance in classification accuracy for each emotion. This was in part driven by the fact that some participants' EEG could not be classified beyond chance levels, while the EEG from others could be classified with over 90% accuracy. We used the results from Task 1 to label those for whom greater than 60% classification accuracy was achieved as 'Classifiable' and those for whom less than 50% accuracy was achieved as 'Non-Classifiable' and trained new classifiers based on these labels. The LOSO-CV classification accuracy for predicting whether participants would be classifiable with respect to each emotion is shown in Figure 1, and the average classifiability across emotions for each participant is shown in Figure 2. These results were driven by temporal theta and beta bicoherence, temporal theta and alpha cross-spectral features, and frontal beta power.

V. Discussion and Conclusions

In this study we showed that it is possible to predict both the presence and magnitude of several different types of emotional experiences to videos with a consumer-grade EEG headband. As such, the method developed here may be more directly applicable to real-world affective computing applications. We achieved average classification accuracies of over 70% when performing LOSO-CV with and without non-classifiable participants (Tasks 1 and 3), LOVO-CV within participants (Task 2), and regular 10-fold CV with and without non-classifiable participants (Tasks 4 and 5). We achieved over 80% classification accuracy with some emotions, such as 'Interest' and 'Anger'. Furthermore, the regression analyses resulted in models which could predict the degree of the reported emotional experience with, in some cases, $R > 0.4$.

For most emotions, we were also able to predict from the EEG which participants would be classifiable or non-classifiable using LOSO-CV. Although the same data were used in the previous classification analyses, a unique set of features were selected when determining whether a participant would be classifiable. This result suggests that it may be possible in the case of emotion recognition or potentially in brain-computer interfacing to determine a priori whether or not a given participants will be easily classifiable. Interestingly, non-classifiability was less of a problem in regression analysis, where the amount of training data seemed to be the main factor in learning accurate models. Further analysis is required to determine why certain participants were non-classifiable. One possibility is that these participants represent a subset of individuals whose emotional reactions produce a different pattern of brain activity from the majority.

Acknowledgments

This research was funded by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to SB and an NSERC PGS scholarship to KD. Thanks to Kristen A. Marszalek for compiling the emotional video database used in this study and for help in data collection. Thanks to Braeden Terpou for help in data collection.

References

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal processing magazine*, vol. 18, no. 1, pp. 32–80, 2001.

[2] J. Tao and T. Tan, "Affective computing: A review," in *International Conference on Affective computing and intelligent interaction*. Springer, 2005, pp. 981–995.

TABLE I: CLASSIFICATION ACCURACY (STD)

	Task 1		Task 2		Task 3		Task 4		Task 5	
	SVM	LR	SVM	LR	SVM	LR	SVM	LR	SVM	LR
Interest	82 (17.6)	73 (27.3)	83 (12.3)	81 (19.2)	80 (14.6)	79 (11.4)	76 (3.0)	76 (3.8)	84 (6.1)	88 (5.6)
Amusement	61 (16.2)	50 (18.3)	71 (13.1)	66 (17.8)	59 (15.5)	47 (14.0)	54 (3.7)	53 (4.5)	59 (6.6)	58 (7.6)
Happiness	61 (13.7)	58 (14.1)	66 (15.0)	59 (12.6)	47 (6.2)	56 (22.3)	55 (2.4)	60 (2.1)	60 (11.5)	58 (15.1)
Sadness	73 (15.6)	70 (19.0)	69 (14.9)	68 (15.8)	76 (5.8)	77 (11.1)	68 (5.8)	70 (3.3)	73 (7.0)	78 (8.1)
Fear	81 (14.9)	79 (17.2)	72 (12.0)	72 (12.0)	80 (5.6)	79 (8.3)	75 (2.8)	77 (2.3)	79 (6.5)	79 (8.0)
Disgust	80 (13.4)	80 (15.3)	75 (8.6)	74 (10.8)	78 (7.9)	77 (10.1)	76 (3.0)	80 (2.5)	78 (1.6)	77 (5.2)
Anger	82 (14.4)	81 (16.7)	78 (10.3)	74 (11.5)	83 (6.8)	83 (9.8)	79 (3.0)	80 (3.1)	84 (2.8)	83 (7.7)
Hope	78 (16.1)	74 (21.9)	78 (12.5)	70 (15.5)	79 (14.2)	79 (11.8)	76 (3.9)	75 (3.6)	80 (6.2)	82 (8.0)
Relief	83 (15.0)	80 (19.1)	76 (12.5)	75 (13.8)	81 (10.6)	75 (22.5)	79 (3.8)	81 (3.3)	76 (4.5)	78 (5.3)
Surprise	73 (19.7)	67 (26.6)	76 (13.2)	71 (16.8)	81 (11.5)	62 (23.9)	68 (3.3)	69 (4.4)	62 (4.6)	61 (6.7)
Sympathy	66 (17.4)	58 (20.5)	73 (13.7)	69 (15.0)	70 (14.6)	66 (12.1)	59 (5.4)	62 (4.1)	71 (6.2)	68 (8.0)
Average	75 (8.4)	70 (10.6)	74 (4.7)	71 (5.6)	74 (11.3)	71 (11.5)	70 (9.5)	71 (9.3)	73 (9.2)	74 (10.6)

TABLE II: REGRESSION MSE and CORRELATION

	Task 1		Task 2		Task 3		Task 4		Task 5	
	MSE	R	MSE	R	MSE	R	MSE	R	MSE	R
Interest	5.17**	0.13*	5.67**	N/A	4.97**	0.28**	4.75**	0.27**	4.89**	0.31**
Amusement	11.31**	0.08	11.50**	0.01	10.53**	0.27**	11.15**	0.12**	11.22**	0.09*
Happiness	13.47**	0.07	21.45	N/A	14.01**	0.25**	13.97**	0.07*	14.95**	0.07
Sadness	12.11**	0.13*	14.88*	0.00	12.68**	0.34**	11.33**	0.13**	11.07**	0.25**
Fear	10.06**	0.09*	8.39**	N/A	7.79**	0.30**	9.34**	0.06*	10.09**	0.15*
Disgust	11.33**	0.06	21.99	N/A	10.26**	0.22**	10.78**	0.08*	11.26**	0.05
Anger	9.29**	0.06	15.37*	0.04	8.65**	0.43**	9.62**	0.02	8.23**	0.34**
Hope	10.02**	0.09	15.09*	0.00	8.44**	0.31**	9.48**	0.11**	9.83**	0.21**
Relief	9.50**	0.03	12.64*	0.04	11.34**	0.15*	9.00**	0.03	5.73**	0.36**
Surprise	11.95**	0.02	11.90*	0.01	6.03*	0.41**	8.72**	0.18**	4.76**	0.38**
Sympathy	11.78**	0.10	13.61*	0.01	11.43*	0.37**	10.26**	0.27**	12.70**	0.25**

* denotes significance from chance level accuracy at $p < 0.05$, and ** denotes significance at $p < 0.005$

[3] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information Fusion*, vol. 37, pp. 98–125, 2017.

[4] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.

[5] C. A. Kothe, S. Makeig, and J. A. Onton, "Emotion recognition from EEG during self-paced emotional imagery," *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*, pp. 855–858, 2013.

[6] D. J. McFarland, M. A. Parvaz, W. A. Sarnacki, R. Z. Goldstein, and J. R. Wolpaw, "Prediction of subjective ratings of emotional pictures by eeg features," *Journal of Neural Engineering*, vol. 14, no. 1, p. 016009, 2016.

[7] S. Makeig, G. Leslie, T. Mullen, D. Sarma, N. Bigdely-Shamlo, and C. Kothe, "First demonstration of a musical emotion BCI," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6975 LNCS, no. PART 2, pp. 487–496, 2011.

[8] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "Eeg-based emotion recognition during watching movies," in *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*. IEEE, 2011, pp. 667–670.

[9] Interaxon, "Muse," www.choosemuse.com, 2014.

[10] MATLAB, *version 8.2.0 (R2013b)*. Natick, Massachusetts: The MathWorks Inc., 2013.

[11] D. H. Brainard, "The psychophysics toolbox," *Spatial vision*, vol. 10, pp. 433–436, 1997.

[12] K. Dhindsa, "Filter-Bank Artifact Rejection: High Performance Real-Time Single-Channel Artifact Detection for EEG," *Biomedical Signal Processing and Control*, In Press: 2017.

[13] J.-P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela *et al.*, "Measuring phase synchrony in brain signals," *Human brain mapping*, vol. 8, no. 4, pp. 194–208, 1999.

[14] U. Greb and M. Rusbridge, "The interpretation of the bispectrum and bicoherence for non-linear interactions of continuous spectra," *Plasma physics and controlled fusion*, vol. 30, no. 5, p. 537, 1988.

[15] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis, and W. V. Friesen, "Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology: I." *Journal of personality and social psychology*, vol. 58, no. 2, p. 330, 1990.

[16] H. C. Peng, "Feature Selection Based on Mutual Information Criteria of Max-dependency, Max-relevance, and Min-redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1226–1238, 2005.

[17] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[18] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005.

[19] T. Hastie, R. Tibshirani, and J. Friedman, "Boosting and additive trees," in *The Elements of Statistical Learning*. Springer, 2009, pp. 337–387.