The Effects of Word Priming on Emotion Classification from Neurological Signals

Cecilia Schmitz
Electrical and Computer Engineering
Kansas State University
Manhattan, KS
clschmitz18@ksu.edu

Taylor Sweet
Electrical and Computer Engineering
Kansas State University
Manhattan, KS
tasweet@ksu.edu

David E. Thompson
Electrical and Computer Engineering
Kansas State University
Manhattan, KS
davet@ksu.edu

Abstract— Affective states play an important role in human behavior and decision-making. In recent years, several affective brain-computer interface (aBCI) studies have focused on developing an emotion classifier based on elicited emotions within the user. However, it is difficult to achieve consistency in elicited emotions across populations, which can lead to dataset imbalances. Observing facial expressions is one method humans use to estimate the internal state of another individual and has been found to elicit more consistent emotions across cultures. The experimental design presented in this paper seeks to avoid consistency issues by asking the participant to classify the emotion portrayed in images of facial expressions, rather than their own emotions. Priming is also a common technique used in psychology studies that is known to influence emotional perception. To improve participant accuracy, we investigated correct and incorrect word priming for the facial expression images. Electro-encephalogram (EEG) data were used to develop a novel classifier that generates an image as an input for a Big Transfer model to classify emotional states. The primed images resulted in higher classification accuracy overall. Further, by building different classifier models for both incorrectly primed images and correctly primed images, we were able to achieve classification accuracies above 90%. We also investigated the effects of providing the classifier with the true labels of the photographs instead of the labels provided by the participants and achieved similar results. The experimental paradigm of measuring brain activity during the emotional classification of another individual provides consistently high, balanced classification accuracies.

Keywords— affective BCI, emotion classification, electroencephalogram, neural network, semantic priming

I. Introduction

Brain-computer interfaces (BCIs) are systems that measure brain activity and often allow brain signals to control an external device [1]. BCIs can also be used to extract specific features from brain signals for analysis of the brain state. In recent decades, emotion classification has become of interest to researchers in many areas of study. Affective brain-computer interfaces (aBCIs) are systems specifically designed for classifying human affect from brain signals [2]. However, despite the popularity of these interfaces in the research community, many design challenges remain. The present work focuses on the creation of reliable emotion elicitation, through the investigation of multiple techniques.

Affect is not a simple phenomenon. Several models have been created for measuring affect. One such model, which examines the link between affective states and facial expressions, is the discrete model of emotions. This model, proposed by Ekman, considers a set of six basic emotions: sadness, happiness, disgust, anger, fear, and surprise [3]. In contrast, J. A. Russel proposed the circumplex model that provides a spectrum of emotions based on two perpendicular axes. The first axis is arousal, which measures the level of activation. The second axis is valence, which measures the level of pleasantness [4-6]. In this work, we will use the discrete model of emotions.

To train and test emotion classification methods, researchers typically seek to elicit a range of emotions in their research participants using stimuli such as images, videos, and audio recordings. Images from the International Affective Picture System (IAPS) are commonly used as visual stimuli because they are standardized according to the arousal and valence emotional dimensions [7]. A second common database used to investigate brain activity and affective states is the DEAP database, a publicly available, multimodal dataset comprised of 32-channel EEG and other physiological signals. The experiment consisted of presenting participants with 40 one-minute long music videos and asking them to rate the video on each axis of the circumplex model [8].

An issue with most, if not all, emotion elicitation paradigms is the difficulty in eliciting the same emotion using the same stimulus on different users. For example, one picture in the IAPS is a gun laying on a table. Although the average population response has been well characterized, guns provoke varying emotions in individuals based on their background. This can

National Science Foundation Grant No. 1910526

lead to a lack of consistency and potentially cause dataset imbalance.

Other studies have used facial expression images because doing so limits the variability across cultures [8]. Following this paradigm, our emotion elicitation paradigm seeks to avoid consistency issues by asking our participants to identify the emotion of the individual in a photograph rather than their own emotions. We used images from the Pictures of Facial Affect (POFA) image set.

In addition to using the POFA, we chose to investigate word priming. Priming is a common psychology technique known to have non-linear effects on emotional perception – as an example, in [9], both matching and mismatched word/picture pairs were shown to participants while measuring emotion recognition time. The response time and accuracy were both improved by matching priming cues, while mismatched priming cues did not adversely affect response time or accuracy relative to no priming. Here, we investigated the effects of word priming on our system accuracy.

II. Methods

A. Participants

Nine (9) participants completed this experiment (6 males and 3 females). All participants were either graduate students or employees of Kansas State University. The range of ages was 29 to 71 with a mean age of 47.4 (STD = 17.6). All participants had normal or corrected-to-normal vision. Only one participant indicated previous use of a BCI. Six (6) participants were right-handed and two (2) were left-handed (one did not specify handedness).

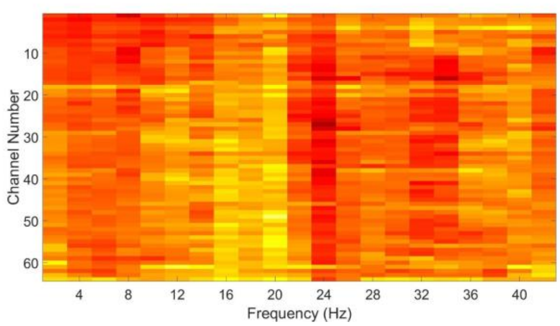

Fig. 2. Heat map of a single trial for a single participant. The real images used as an input for the BiT-MR101x1 model did not include the axis labels.

B. Experimental Setup

This experiment used 60 gray scale images from the Pictures of Facial Affect (POFA) image set. Each image consists of an actor expressing one of the following emotional states: happy, sad, anger, disgust, fear, and surprise. In previous experiments, the images were ranked by the percent of participants who correctly identified the portrayed emotion [10]. The ten (10) images with the lowest percent of each emotion class were selected for this study. A set of word primes was also created for each of the emotional states. The images were presented on a 15-inch PC monitor with a refresh rate of 60Hz and resolution of 1920 x 1080. Participants were seated approximately 2 feet away from the monitor. A 64-channel Cognioncs Mobile 72 EEG was placed according to the international 10-20 system and used to record signals at a sampling rate of 500 Hz. ABRALYT High-Chloride gel was used to reduce the impedance below 50 kΩ.

C. Experimental Design

The experiment involved two sessions of 1 hour, including time for setting up the BCI. The sessions were further split into five separate runs to limit participant fatigue. Each run consisted of a BCI2000 stimulus presentation task using 12 facial images, totaling 60 images for the session. Alternating images were preceded by a word prime, of which half correctly matched the emotional state of the image. The images selected to be primed were chosen randomly. The word primes were displayed for one second and the facial images were displayed for five seconds. After each facial image, a 10 second blank image was shown to allow the participant time to identify the emotion portraying in the facial image on a response sheet. The participant was also asked to provide an estimate of their confidence in their ability to correctly identify the emotion. The confidence rating was on a scale of one to five. The design of the stimuli presentation pattern is displayed in Fig. 1 below. In the second session, the same 60 images were used; however, the primed and unprimed images from the first session were swapped.

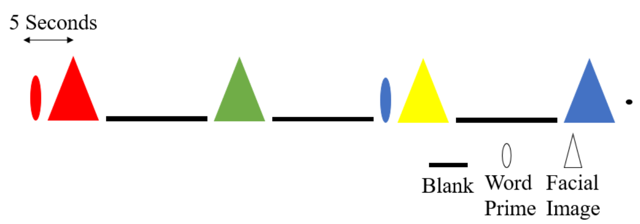

Fig. 1. Graphical representation of stimuli presentation. The ovals represent the word primes, the triangles represent the images, and the lines represent the blanks. The different colors are distinct emotions. When the oval matches in color with the succeeding triangle, the image is correctly primed. When the triangle is not preceded by an oval, the image is unprimed. The figure is made to scale to represent 1 second for the word prime, 5 seconds for the image, and 10 seconds for the blank. The pattern in the figure repeats three (3) times during a run.

D. Data Preprocessing

The EEG signal was filtered using an 8th order lowpass infinite impulse response (IIR) filter with a cutoff frequency of 45 Hz. The signal was then split into one trial per facial image using a 0 to 900 millisecond window, where the initial timepoint corresponds to the stimulus onset. The 500 milliseconds preceding the word prime, if present, or facial image was selected as the baseline for the trial. The Fourier transform was applied to both the baseline and the signal for each channel, trial, and participant. The resulting frequency coefficients for the baseline and signal were separated into bins with a size of 2 Hz, ranging from 2 Hz to 44 Hz. Finally, the coefficients in each bin were summed and the values were used to convert to decibels. A heat map representation of channel by frequency was created for each trial for each participant to use as an input for the BiT-MR101x1 model [11] and is shown in Fig. 2 below.

TABLE I. CLASSIFIER ACCURACY USING PARTICIPANT LABELS

	Participant Labels		
	Accuracy	*Balanced Accuracy*	*Images per Participant*
Unprimed	64.81%	63.25%	60
Primed	66.67%	64.80%	60
Correctly Primed	92.59%	89.63%	30
Incorrectly Primed	92.59%	92.78%	30

E. Emotion Classification

After the heat maps were generated for each trial, they were resized to 512 by 512 pixels and the color values were normalized. Images of the heat maps were taken and divided to use 70% for the training set and 30% for the test set. The images were then input into the BiT-M R101x1 model, which was used as the classifier. A separate classifier was trained over 30 epochs, with 10 steps per epoch, for each participant. The neural network was set to use a Stochastic Gradient Descent Optimizer and a Sparse Categorical Cross entropy loss function. The high-performance computing cluster, Beocat, at Kansas State University was used to decrease the training time by allowing for the training of multiple classifiers simultaneously.

TABLE II. CLASSIFIER ACCURACY USING PARTICIPANT LABELS

	Image Labels		
	Accuracy	*Balanced Accuracy*	*Images per Participant*
Unprimed	71.60%	73.83%	60
Primed	71.53%	71.87%	60
Correctly Primed	92.59%	95.12%	30
Incorrectly Primed	88.89%	88.70%	30

III. RESULTS

In this experiment, we introduced word priming to measure the effects on our ability to classify the emotion portrayed in the image based on participant EEG data. The EEG recordings from each trial were used to estimate the affective state of the participants in response to attempting to identify the emotion in the photograph. Table I displays the emotion classification accuracies of the different stimuli configurations based on the labels provided by the participants. The primed images resulted in a higher emotion classification accuracy than the unprimed images. The incorrectly primed images also had a slightly higher balanced accuracy than the correctly primed. The images per participant used to for each stimuli configuration are also notable because fewer images could be used to train each of the correctly and incorrectly primed classifiers.

The EEG recordings from each trial were also used to estimate the affective state based on the true labels of the facial images. The classifier accuracies for each of the stimuli configurations are included in Table II. For the primed and unprimed images, the accuracies were higher using the image labels than using the participant labels. The correctly primed images also had higher accuracies, while the incorrectly primed had lower accuracies relative to Table I. Using the image labels, the difference between the correctly and incorrectly primed classification accuracy is larger.

In Fig. 3. we plotted the accuracy values for all classifiers. All classifiers showed significantly higher accuracy than chance. We also see no significant difference between the unprimed and primed classifiers using either labeling scheme. However, both the incorrectly primed and correctly primed classifiers yielded significantly higher accuracy than the unprimed classifier, regardless of the labeling scheme.

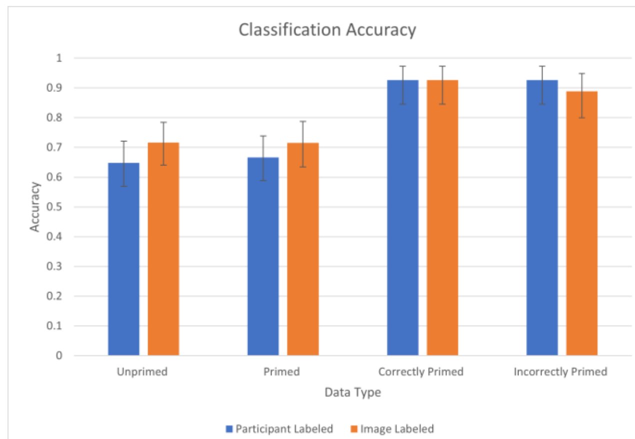

Fig. 3. Emotion classification accuracies based on participant labels (blue) and true image labels (orange). The confidence bounds for each accuracy per each stimuli configuration are displayed.

IV. CONCLUSION

Overall, the experiment resulted in reliable accuracy values; the classification accuracy of all image configurations was significantly higher than chance accuracy. While we saw no significant difference between the primed and unprimed classifiers, we were able to achieve significantly higher accuracies when we further separated the primed images into incorrectly primed and correctly primed.

Although the results are promising, the experiment has several limitations. Typically, aBCI studies seek to elicit emotion and record the corresponding physiological signals. Our study did not seek to elicit an emotional state and, instead, had participants estimate the emotional state of a person in a photograph. Therefore, we may not be performing emotional classification of the participant. We estimate that 20 participants would have 92% of the power to detect a 10% difference in accuracy; however, our study currently includes nine (9) participants. This study is ongoing, and we plan to continue until we reach 20 participants.

V. Discussion

Most common aBCI research involves eliciting emotions in the research participants for affective state classification. However, due to the complexity of affect and variability in elicited responses, this experimental paradigm can lead to inconsistency issues, such as dataset imbalance. The present work focuses on the creation of reliable emotion elicitation, through the investigation of multiple techniques. One of which is a new experimental paradigm where we asked participants to identify the emotion portrayed in photographs of facial expressions using the six (6) emotions from the discrete model of emotions. The second is preceding alternating images with word primes, which would either match the emotional state of the image or be the word of one of the other five (5) emotional states. During this process, we collected EEG signals from 64 channels to create image representations of their neurological state and an image classifier to determine which of the six (6) emotions the participant selected. We found that, if we built different models for both incorrectly primed images and correctly primed images, we were able to achieve accuracies above 90%. We also investigated the effects of providing the classifier with the true labels of the photographs instead of the labels provided by the participants and achieved very similar results.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1910526. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work was partially funded by the Women in Engineering program at Kansas State University. The computing for this project was performed on the Beocat Research Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860, EPS-0919443, ACI-1440548, CHE-1726332, and NIH P20GM113109.

References

[1] J. R. Wolpaw et al., "Brain-computer interface technology: a review of the first international meeting," IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 2, pp. 164–173, Jun. 2000, doi: 10.1109/tre.2000.847807.

[2] M. R. Mowla, R. I. Cano, K. J. Dhuyvetter, and D. E. Thompson, "Affective Brain-Computer Interfaces: Choosing a Meaningful Performance Measuring Metric," Computers in Biology and Medicine, vol. 126, p. 104001, Nov. 2020, doi: 10.1016/j.compbiomed.2020.104001.

[3] P. Ekman, E. R. Sorenson, and W. V. Friesen, "Pan-Cultural Elements in Facial Displays of Emotion," Science, vol. 164, no. 3875, pp. 86–88, Apr. 1969, doi: 10.1126/science.164.3875.86.

[4] L. Camras and R. Plutchik, "Emotion: A Psychoevolutionary Synthesis," The American Journal of Psychology, vol. 93, no. 4, p. 751, Dec. 1980, doi: 10.2307/1422394.

[5] C. M. Whissell, "The Dictionary of Affect in Language," The Measurement of Emotions, pp. 113–131, 1989, doi: 10.1016/b978-0-12-558704-4.50011-6.

[6] J. A. Russell, "A circumplex model of affect.," Journal of Personality and Social Psychology, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.

[7] G. Liberati, S. Federici, and E. Pasqualotto, "Extracting neurophysiological signals reflecting users' emotional and affective responses to BCI use: A systematic literature review," NeuroRehabilitation, vol. 37, no. 3, pp. 341–358, Oct. 2015, doi: 10.3233/nre-151266.

[8] S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: 10.1109/t-affc.2011.15.

[9] N. C. Carroll and A. W. Young, "Priming of emotion recognition," Q. J. Exp. Psychol. Sect. A, vol. 58, no. 7, pp. 1173–1197, Oct. 2005.

[10] A. Kolesnikov et al., "Big Transfer (BiT): General Visual Representation Learning," arXiv:1912.11370 [cs], May 2020, [Online]. Available: https://arxiv.org/abs/1912.11370.

[11] P. Ekman and W. V. Friesen (1976). Pictures of facial affect. Palo Alto, CA: Consulting Psychological Press.