Automated Detection of Optimal DBS Device Settings

Yaohan Ding
yad30@pitt.edu
University of Pittsburgh
Pittsburgh, U.S.A.

Itir Onal Ertugrul
i.onal@tilburguniversity.edu
Tilburg University
Tilburg, Netherlands

Ali Darzi
ald260@pitt.edu
University of Pittsburgh
Pittsburgh, U.S.A.

Nicole Provenza
nicole_provenza@brown.edu
Brown University
Providence, U.S.A.

László A. Jeni
laszlojeni@cmu.edu
Carnegie Mellon University
Pittsburgh, U.S.A.

David Borton
david_borton@brown.edu
Brown University
Providence, U.S.A.

Wayne Goodman
wayne.goodman@bcm.edu
Baylor College of Medicine
Houston, U.S.A.

Jeffrey Cohn*
jeffcohn@pitt.edu
University of Pittsburgh
Pittsburgh, U.S.A.

ABSTRACT

Continuous deep brain stimulation (DBS) of the ventral striatum (VS) is an effective treatment for severe, treatment-refractory obsessive-compulsive disorder (OCD). Optimal parameter settings are signaled by a mirth response of intense positive affect, which is subjectively identified by clinicians. Subjective judgments are idiosyncratic and difficult to standardize. To objectively measure mirth responses, we used Automatic Facial Affect Recognition (AFAR) in a series of longitudinal assessments of a patient treated with DBS. Pre- and post-adjustment DBS were compared using both statistical and machine learning approaches. Positive affect was significantly higher after DBS adjustment. Using XGBoost and SVM, the participant's pre- and post-adjustment responses were differentiated with accuracy values of 0.76 and 0.75, which suggest feasibility of objective measurement of mirth response.

CCS CONCEPTS

• **Applied computing** → **Psychology**.

KEYWORDS

affective computing; clinical research; deep brain stimulation (DBS); ventral striatum; obsessive-compulsive disorder (OCD)

ACM Reference Format:
Yaohan Ding, Itir Onal Ertugrul, Ali Darzi, Nicole Provenza, László A. Jeni, David Borton, Wayne Goodman, and Jeffrey Cohn*. 2020. Automated Detection of Optimal DBS Device Settings. In *Companion Publication of the 2020 International Conference on Multimodal Interaction (ICMI '20 Companion), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3395035.3425354

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-8002-7/20/10.
https://doi.org/10.1145/3395035.3425354

1 INTRODUCTION

Obsessive compulsive disorder (OCD) is a persistent, oftentimes disabling psychiatric disorder that is characterized by obsessive thoughts and compulsive behavior. Obsessive thoughts are intrusive and unwanted and can be highly disturbing. Compulsions are repetitive behaviors that an individual feels driven to perform. For patients with severe treatment-resistant OCD, continuous deep brain stimulation (DBS) of the ventral striatum (VS) is an effective treatment [1]. Nearly 70 percent of patients experienced a 25 percent or greater decrease in symptom severity[10] after a 12-month treatment[5]. While the underlying principles and mechanisms of DBS are not fully understood, DBS directly changes the activation of the target region in a controlled manner [8]. The target region is the ventral striatum, which is involved in affect processing. A potential side effect of DBS is hypomania, which can have deleterious consequences. To avoid this potential side effect and maximize treatment efficacy, optimal programming of DBS is essential over the course of treatment.

A mirth response of intense positive affect frequently occurs during initial DBS and signals good prognosis. The mirth response is related to the affective circuitry of the VS. In practice, optimal DBS adjustments are made largely on the basis of subjective clinical judgment of patient responses. While useful, subjective judgments are idiosyncratic and difficult to standardize. To maximize treatment efficacy while minimizing potential side effects, we applied objective, automatic measurement to programming session videos.

The mirth response is quantified as the intensity summation of two facial actions units (AU) as defined by the Facial Action Coding System [7]: AU6 (cheek raiser) and AU12 (lip corner puller), which together make the Duchenne smile that represents and signals positive emotion[2, 3, 9, 11]. We acquired the intensity of these 2 AUs using Automatic Facial Affect Recognition (AFAR) [14], a powerful toolkit for assessing severity of negative and positive affect[6, 12, 13]. We explored its effectiveness to objectively measure mirth response (positive affect) to DBS adjustments.

2 METHODS AND EXPERIMENTS

2.1 A. Data and Features

This case study is from an ongoing clinical trial for treatment-refractory OCD. The participant, a male, is the first one to complete the two-year long trial. Over this period, DBS was manually adjusted by his physician in a series of 21 video-recorded programming sessions beginning at his sixth session. In prior sessions (1 through 5) the DBS device had not yet been implanted or had not yet been activated with brief exception of during implantation surgery.

In 11 sessions, DBS was adjusted to an optimal setting from its initial setting for that session. In some others (16, 21, 22, 25 and 26) no change was made or change was not recorded (17 and 18); and three sessions were missing for lack of video (20) , an incomplete visit (23) and technical reasons (24). Our primary interest was in detecting difference in positive affect between pre- and post-adjustment settings in the 11 sessions in which adjustments in DBS occurred.

For each session, AFAR tracks 49 2D-facial landmarks, 3 head poses (pitch, yaw, and roll), and intensities of facial action units . Positive affect (PA) in facial expression was represented by Duchenne smiling: adding up the estimated intensities of AU 6 (cheek raiser) and AU 12 (lip corner puller). Frame-level displacement and velocity of head pose and facial muscle movement (averaging over 49 landmarks) were calculated to represent the dynamics of head and face. Amplitude and velocity of AU6 and AU12 intensities were calculated to represent the dynamics of AUs.

Because DBS targets affective circuits, a mirth response is expected to occur rapidly. For this reason, we focused on the first 15s trial of the pre- and optimal post-adjustment settings in each of the 11 sessions provided by the clinical record. Some minor exceptions were necessary. In Session 7, because of large head motion that confounded initial face tracking, we shifted to the first tracked 15s trial. Session 13's initial setting lasted only 4 seconds. Session 14's final setting lasted 12s.

For machine learning model training, each 15s trial was divided into 15 1-second samples and assigned to 3 session-independent folds. For each sample, we calculated the mean, max, and standard deviation of the 8 frame-level features described above, resulting in a 1x24 feature vector for each sample.

XGBoost[4] and SVM[15] classifiers were trained to distinguish participant's responses under pre- and optimal post-adjustment DBS settings. For XGBoost, features were later selected based on the summation of their weights in 3 folds. We sorted the features by weight descending and built new models by adding one feature each time until the performance peaked. For SVM, we used a sequential forward feature selection strategy where the top 24 updated feature sets were retained for the next feature-adding round. The final feature sets were selected based on the best performance and the least number of features. Because of the limited number of samples and our interest in features that can be interpreted, deep approaches were not considered. Since the dataset is balanced, classifier performance is evaluated by accuracy. The chance accuracy is calculated by assuming the models only predict the post-adjustment class.

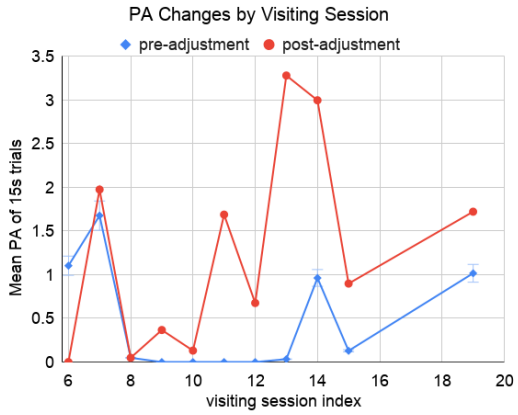

Figure 1: Change in PA by visiting session. Programming started at the 6th visit.

Table 1: 3-fold session-independent cross validation classification results

Trial Length	#Sample	Chance Accuracy	XGBoost #Selected Feature	XGBoost Accuracy	SVM #Selected Feature	SVM Accuracy
3	66	0.50	1	0.64	3	0.74
5	109	0.50	1	0.68	9	0.72
7	151	0.51	2	**0.76**	9	**0.75**
10	214	0.51	2	0.75	11	0.70
12	256	0.52	3	0.70	6	0.69
15	316	0.51	8	0.64	12	0.71

2.2 B. Results

(**Figure. 1**) shows the mean PA of 15s trials under two settings by session indices. The studied sessions are marked. For 9 out of 11 sessions, the participant evidenced increases in positive affect (PA) with $t = -2.29,\ p < 0.05$ (paired t-test with 10 $degrees\ of\ freedom$). We also examined intervals over the range 3 to 15. All were statistically significant.

Accuracy for the 15s trials and for shorter ones (**Table.1**) suggests that both classifiers perfomred better than chance. Seven seconds was the best window length for both XGBoost and SVM classifiers to differentiate pre- and post-adjustment responses. Classifiers trained for this purpose are a promising tool for clinical decision making.

While the two approaches yielded similar results, they differed in the number of features used and in consistency of their findings. XGBoost used fewer features at the cost of slightly increased variability across trial lengths. While SVM accuracies ranged from 0.69 to 0.75, XGBoost accuracies ranged from 0.64 to 0.76. In terms of feature selection in 6 trials, XGBoost models always assigned high weights to the mean AU12 intensity, mean facial displacement and mean head displacement, whereas SVM always retained highly correlated features (mean AU6 velocity, mean AU12 intensity and mean AU12 velocity) in final feature sets. In the future, we are interested to see whether this pattern is repeated in additional participants and look for better feature representations.

3 ACKNOWLEDGEMENTS

This work was supported in part by NIH award UH3NS100549.

REFERENCES

[1] Pino Alonso, Daniel Cuadras, Loes Gabriëls, Damiaan Denys, Wayne Goodman, Ben D Greenberg, Fiacro Jimenez-Ponce, Jens Kuhn, Doris Lenartz, Luc Mallet, et al. 2015. Deep brain stimulation for obsessive-compulsive disorder: a meta-analysis of treatment outcome and predictors of response. *PloS one* 10, 7 (2015), e0133591.

[2] Zara Ambadar, Jeffrey F. Cohn, and Lawrence Ian Reed. 2009. All Smiles are Not Created Equal: Morphology and Timing of Smiles Perceived as Amused, Polite, and Embarrassed/Nervous. *Journal of Nonverbal Behavior* 33, 1 (mar 2009), 17–34. https://doi.org/10.1007/s10919-008-0059-5

[3] Jason K. Baker, John D. Haltigan, Ryan Brewster, James Jaccard, and Daniel Messinger. 2010. Non-expert ratings of infant and parent emotion: Concordance with expert coding and relevance to early autism risk. *International Journal of Behavioral Development* 34, 1 (jan 2010), 88–95. https://doi.org/10.1177/0165025409350365

[4] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 785–794.

[5] Damiaan Denys, Ilse Graat, Roel Mocking, Pelle de Koning, Nienke Vulink, Martijn Figee, Pieter Ooms, Mariska Mantione, Pepijn van den Munckhof, and Rick Schuurman. 2020. Efficacy of deep brain stimulation of the ventral anterior limb of the internal capsule for refractory obsessive-compulsive disorder: a clinical cohort of 70 patients. *American Journal of Psychiatry* 177, 3 (2020), 265–271.

[6] Hamdi Dibeklioğlu, Zakia Hammal, and Jeffrey F Cohn. 2017. Dynamic multimodal measurement of depression severity using deep autoencoding. *IEEE journal of biomedical and health informatics* 22, 2 (2017), 525–536.

[7] Paul Ekman, W. V. Friesen, and J. C. Hager. 2002. Facial action coding system: Research Nexus. In *Network Research Information, Salt Lake City, UT.*

[8] Míriam R García, Barak A Pearlmutter, Peter E Wellstead, and Richard H Middleton. 2013. A slow axon antidromic blockade hypothesis for tremor reduction via deep brain stimulation. *PloS one* 8, 9 (2013), e73456.

[9] Jeffrey M. Girard, Jeffrey F. Cohn, Lijun Yin, and Louis-Philippe Morency. 2020. Reconsidering the Duchenne smile: Examining the relationships between the Duchenne marker, smile intensity, and positive emotion. *PsyArXiv. doi:10.31234/osf.io/397af* (2020).

[10] Wayne K Goodman, Lawrence H Price, Steven A Rasmussen, Carolyn Mazure, Roberta L Fleischmann, Candy L Hill, George R Heninger, and Dennis S Charney. 1989. The Yale-Brown obsessive compulsive scale: I. Development, use, and reliability. *Archives of general psychiatry* 46, 11 (1989), 1006–1011.

[11] Nathaniel Haines, Matthew W. Southward, Jennifer S. Cheavens, Theodore Beauchaine, and Woo-Young Ahn. 2019. Correction: Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity. *PLOS ONE* 14, 3 (mar 2019), e0213756. https://doi.org/10.1371/journal.pone.0213756

[12] Zakia Hammal, Jeffrey F Cohn, and David T George. 2014. Interpersonal coordination of headmotion in distressed couples. *IEEE transactions on affective computing* 5, 2 (2014), 155–167.

[13] Katherine B Martin, Zakia Hammal, Gang Ren, Jeffrey F Cohn, Justine Cassell, Mitsunori Ogihara, Jennifer C Britton, Anibal Gutierrez, and Daniel S Messinger. 2018. Objective measurement of head movement differences in children with and without autism spectrum disorder. *Molecular autism* 9, 1 (2018), 14.

[14] I. Onal Ertugrul, L. A. Jeni, W. Ding, and et al. 2019. AFAR: A Deep Learning Based Tool for Automated Facial Affect Recognition. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*. IEEE. https://github.com/AffectAnalysisGroup/AFARtoolbox

[15] Johan AK Suykens and Joos Vandewalle. 1999. Least squares support vector machine classifiers. *Neural processing letters* 9, 3 (1999), 293–300.