Automatic Detection of Target Engagement in Transcutaneous Cervical Vagal Nerve Stimulation for Traumatic Stress Triggers

Nil Z. Gurel, *Student Member, IEEE*, Matthew T. Wittbrodt, Hewon Jung, Stacy L. Ladd, Amit J. Shah, Viola Vaccarino, J. Douglas Bremner, and Omer T. Inan, Senior *Member, IEEE*

***Abstract*—Transcutaneous cervical vagal nerve stimulation (tcVNS) devices are attractive alternatives to surgical implants, and can be applied for a number of conditions in ambulatory settings, including stress-related neuropsychiatric disorders. Transferring tcVNS technologies to at-home settings brings challenges associated with the assessment of therapy response. The ability to accurately detect whether tcVNS has been effectively delivered in a remote setting such as the home has never been investigated. We designed and conducted a study in which 12 human subjects received active tcVNS and 14 received sham stimulation in tandem with traumatic stress, and measured continuous cardiopulmonary signals including the electrocardiogram (ECG), photoplethysmogram (PPG), seismocardiogram (SCG), and respiratory effort (RSP). We extracted physiological parameters related to autonomic nervous system activity, and created a feature set from these parameters to: 1) detect active (vs. sham) tcVNS stimulation presence with machine learning methods, and 2) determine which sensing modalities and features provide the most salient markers of tcVNS-based changes in physiological signals. Heart rate (ECG), vasomotor activity (PPG), and pulse arrival time (ECG+PPG) provided sufficient information to determine target engagement (compared to sham) in addition to other combinations of sensors. resulting in 96% accuracy, precision, and recall with a receiver operator characteristics area of 0.96. Two commonly utilized sensing modalities (ECG and PPG) that are suitable for home use can provide useful information on therapy response for tcVNS. The methods presented herein could be deployed in wearable devices to quantify adherence for at-home use of tcVNS technologies.**

***Index Terms*—traumatic stress, physiological biomarkers, vagal nerve stimulation, wearable bioelectronic medicine, wearable neuromodulation, wearable sensing.**

I. Introduction

THE vagus nerve is a complex cranial nerve which, via connections to the brain, neck, heart, lungs, and abdomen, mediates autonomic tone within the body. Efferent vagal projections interface with peripheral organs to mediate autonomic, endocrine, and behavioral responses [1]. Electrical stimulation of the vagus nerve has promising therapeutic effects on inflammatory, psychiatric, and cardiovascular disorders, either through implantable or noninvasive devices [2, 3]. The latter method includes transcutaneous stimulation of the auricular branch (in the ear) or the cervical branch (at the neck). Auricular and cervical VNS (taVNS and tcVNS) appear to activate cortical brain areas receiving vagal projections [4] along with limbic structures [5] which have been hypothesized to affect baroreception and priopriocepion. Noninvasive vagal stimulation technologies have similar therapeutic benefits compared to the implanted counterparts, while avoiding the need for surgery. These noninvasive technologies have great potential for daily at-home use for rehabilitation, mood, and performance improvement [6, 7]. Moreover, existing research demonstrates their safety and tolerability for use in human subjects [8, 9].

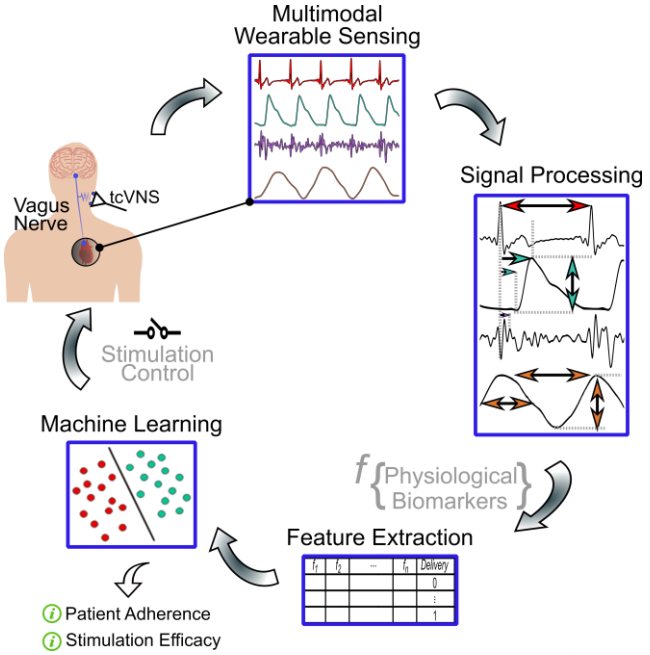

Fig. 1. Wearable neuromodulation technologies can interface with noninvasive sensing. Ongoing activity acquired from one or more sensors can be utilized to assess cardiovascular and peripheral physiological response via signal processing. Events or changes in these data streams can then be decoded using machine learning via feature extraction, in order to dynamically trigger closed-loop delivery when needed, and to quantify patient adherence and stimulation efficacy for at-home therapy.

Although taVNS and tcVNS have similar effects in imaging studies, they should be considered separate as the stimulation targets different locations of the vagus nerve -auricular branch through the ear and cervical branch through the neck, respectively. Hour-long sessions of taVNS have been shown to improve vagal tone [10, 11]. Favorable outcomes have been reported for migraine [12], epilepsy [13], major depression [14, 15], posttraumatic stress disorder [16]. Pairing taVNS with

acute psychological stress was also studied: taVNS was shown to improve subjective fear and worry responses [17-19]. The newer and less common approach, tcVNS, was shown to decrease serum cytokines, and improve cardiac vagal tone [20, 21]. Clinically-relevant effects of minute-long sessions of tcVNS have been reported for conditions such as trigeminal allodyna [22], and cluster headache [23]. tcVNS is particularly suitable for acute therapies due to short session time (two minutes) when stress-related neuropathways are activated.

The investigation of tcVNS applied in tandem with acute traumatic stress is clinically relevant. Patients with trauma-related psychiatric disorders may experience traumatic flashbacks multiple times a day, triggered by, for example, sensory data (odor, sound, room size), cues related to the traumatic event, or even temporal data (incident repeated at the same time of the day) [24-28]. These specific triggers may be relevant to personal traumatic memories from sexual assault to combat exposure [29]. Recently, we demonstrated clinically-relevant physiological effects of tcVNS as measured by downstream physiological signals [30], as a first step to evaluate the merit of tcVNS usage in context of traumatic stress. Engineering efforts to transfer tcVNS to at-home use bring challenges regarding the stimulation application. A fundamental challenge is the determination of whether stimulation delivered properly [31]: The electrode placement and stimulation intensity affect the current flow and distribution, hence the stimulation efficacy, as tcVNS electrodes are directly in contact with the skin rather than placed on the nerve itself (as with an implantable counterpart). *Thus, personalized neuromodulation to provide feedback for the physician or for the patient on whether stimulation has been properly delivered to the vagus nerve is necessary for longitudinal treatment paradigms.* As summarized in Fig. 1, interfacing tcVNS with noninvasive sensing modalities and extracting information to decode whether the stimulation engaged the nerve target could dramatically improve stimulation efficacy and provide closed-loop delivery in response to a detected event. Previous work from our group has demonstrated that certain physiological parameters appear to be reflective of tcVNS delivery compared to sham [30, 32, 33], but the ability to accurately detect whether tcVNS has been delivered has never been investigated. Continuous cardiovascular and peripheral physiological sensing offer a convenient tool for this investigation, due to the intimate relationship of the vagus with the heart and the peripheral physiology, and the ubiquity of noninvasive wearable technologies. Previously, changes in individual features such as heart rate (HR) have been observed, the differences between groups in these individual features were not sufficient to allow for classifying whether or not tcVNS was properly delivered. Another point to consider is the optimization of the acquired modalities: while wearable sensing offers a convenient tool to quantify tcVNS, learning from the optimized modalities (i.e. "learning from less data") would be favorable for hardware adaptations that possibly may not employ all measurement modalities discussed.

In this work, we conducted a study with 26 human subjects, where 12 received active tcVNS and 14 received sham stimulation in tandem with acute traumatic stress. We collected continuous physiological signals related to the electrical and mechanical activity of the heart, peripheral blood volume, and respiration from electrocardiography (ECG), seismocardiography (SCG), photoplethysmography (PPG), respiratory effort (RSP) signals. Utilizing signal processing methods, we extracted parameters that are indicative of autonomic nervous system (ANS) activity, peripheral vasoconstriction, and respiration dynamics. Then, we employed machine learning techniques to determine the stimulation type, active or sham, from the extracted parameters and determined suitable sensing modalities for "in the wild" tcVNS use that could quantify therapy response, and ultimately enable closed-loop tcVNS delivery when a traumatic stress trigger is detected.

TABLE I
DEMOGRAPHICS AND BASELINE PHYSIOLOGICAL PARAMETERS

Parameter	Active (n=12)	Sham (n=14)	P-value
Age [years]	28.7 (6.8)	30.8 (10.6)	0.57
Sex [F, %]	5, 41.7%	9, 64.2%	0.25
Weight [kg]	76.8 (13.9)	78.8 (14.1)	0.86
Height [cm]	174.7 (11.1)	170.5 (6.7)	0.47
BMI [kg/m^2]	25 (3.2)	26.9 (4.2)	0.37
HR [bpm]	67.4 (13.9)	62.4 (9)	0.63
PEP [ms]	66.5 (17.7)	62.4 (9)	0.98
PPG Amplitude [a.u.]	0.3 (0.3)	0.1 (0.1)	0.11
PAT_{FOOT} [ms]	274 (31.7)	261.2 (19.2)	0.34
PAT_{PEAK} [ms]	450.6 (38.2)	437.1 (34.1)	0.34
RR [breath-pm]	17.3 (5.4)	18. (3)	0.63
RW [s]	1.8 (0.7)	1.6 (0.3)	0.66
RP [a.u.]	0.8 (0.6)	1.3 (0.9)	0.27

F = female; BMI = Body-mass index; bpm = beats per minute; breath-pm = breaths per minute, values indicate mean (SD).

II. METHODS

A. Human Subjects Study and tcVNS Application

The study focused on tcVNS effects on acute traumatic stress and was conducted under a protocol approved by the institutional review boards (IRB) of the Georgia Institute of Technology (#H17126), Emory University (#IRB00091171), SPAWAR Systems Center Pacific, and the Department of Navy Human Research Protection Office (HRPO), between May 2017 to October 2019 (ClinicalTrials.gov # NCT02992899). A total of 26 adults aged 18-65 who had experienced at least one psychologically traumatic event and based on Diagnostic and Statistical Manual-5 (DSM-5) criteria were recruited [34, 35]. Personal traumatic events were recorded upon recruitment and presented as voice recordings during the study session. After a baseline resting period, the stimulation was applied to the left side of the neck immediately after hearing the one-minute traumatic stress recording. The subjects kept supine position throughout the protocol. The protocol included two kinds of

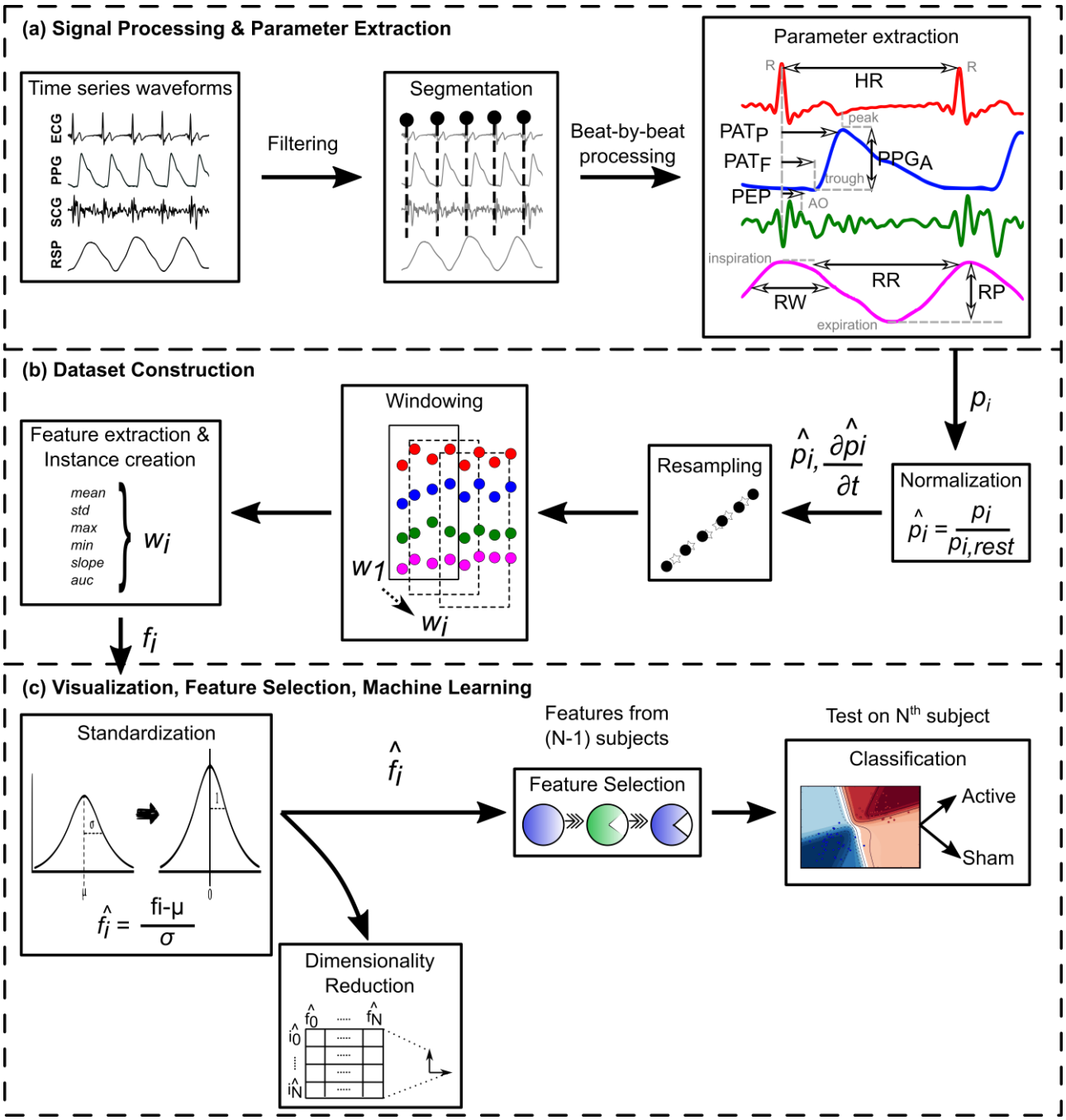

Fig. 2. (a) ECG, PPG, SCG, RSP signals were processed and HR, PAT, PEP, PPG amplitude, RR, RW, RP were extracted as physiological parameters. (b) Using the extracted parameters, dataset constructed after normalization, resampling, and windowing. (c) After standardization, dimensionality reduction methods were applied for dataset visualization. Then, feature selection and machine learning were conducted. PAT_F: PAT_{FOOT}; PAT_P: PAT_{PEAK}; AO: Aortic opening; PPG_A: PPG amplitude, μ: mean; σ: standard deviation; p: parameters; w: windows; f: features.

stimuli: active tcVNS or sham stimulus (GammaCore, ElectroCore, Basking Ridge, NJ), both of which were indistinguishable in appearance and operation with the following waveform characteristics: 5kHz sine bursts repeating at a rate of 25 Hz for the active tcVNS (0-30V peak output, adjustable intensity), and 0.2 Hz stepped pulse (0-14V peak output, adjustable intensity) for the sham stimulus, both lasting for two minutes. The subjects, the researchers, and the clinical staff were blinded to the device type. As the vagus nerve is typically located by the carotid artery, collar electrodes were placed on the left neck on top of the carotid pulse using a conductive electrode gel (GammaCore, ElectroCore, Basking Ridge, NJ) and the devices were operated by the researcher.

B. *Statistical Analysis*

Subject demographics (age, gender, height, weight, body-mass index) and baseline physiological parameters were compared between the two device groups, active and sham, to understand whether the groups were significantly different from each other. Normality was assessed using Shapiro-Wilk test. The comparisons were made using t-tests for normal continuous variables, Wilcoxon rank-sum tests for non-normal continuous variables, and chi-squared tests for categorical variables. P-

values lower than 0.05 were considered statistically significant. Table I summarizes the demographics and baseline physiological parameter information for the subjects, for each device group: active and sham. The groups were determined to be balanced with equivalent baseline characteristics, rendering the data amenable to use for further engineering purposes.

C. Physiological Monitoring

ECG, SCG, PPG, RSP signals were collected continuously throughout the protocol using wireless amplifiers for 3-lead ECG, RSP, and transmissive PPG signals (Bionomadix RSPEC-R and PPGED-R, Biopac Systems, Goleta, CA). SCG was measured in the dorsoventral direction using a low noise accelerometer placed on the middle of the sternum (356A32, PCB Electronics, Depew, NY). RSP measurement was taken using a respiration belt and the PPG measurement was taken from the index finger. All data were simultaneously recorded using a 16-bit data acquisition system at 2kHz (MP150, Biopac Systems, Goleta, CA).

D. Signal Processing and Parameter Extraction

Fig. 2a summarizes signal processing and parameter extraction paradigm, conducted in MATLAB (R2017b, Natick, MA). The physiological parameters related to ANS and peripheral physiological activity were extracted from four sensing modalities: ECG, PPG, SCG, RSP. First, the signals ECG, PPG, SCG were band-pass filtered with finite impulse response filters with cut-off frequencies 0.6-40Hz for ECG, 0.4-8Hz for PPG and 0.5-25Hz for SCG [36]. PPG and SCG signals were beat-by-beat segmented with beat lengths of 150ms and 600ms, respectively, with reference to the R-peaks of the ECG signal that are located by thresholding. These beat lengths were selected as they were adequate to capture the fiducial points in each beat. Exponential moving averaging was implemented for some of the parameter extraction tasks listed below to reduce motion artifacts [37]. The extracted parameters are listed below.

Heart rate (HR): HR was extracted for multiple reasons. First, it is a basic vital sign used in clinical studies. Second, and more importantly, HR is regulated by both branches of the vagus-regulated ANS, the sympathetic (SNS) and parasympathetic nervous systems (PNS) [38]. Multiple implantable (direct) and noninvasive VNS studies noted changes in HR [23, 39, 40]. Therefore, instantaneous HR was computed from the R-R intervals of the ECG signals (in beats per minute) as a complex indicator of ANS balance.

PPG amplitude and pulse arrival time (PAT): The PPG signal is a rich and complex source of information about arterial tone (vasodilation or vasoconstriction), peripheral sympathetic activity, and relative changes in pulse pressure [41]; therefore, multiple parameters were extracted from this signal. The beat-by-beat amplitude of the PPG waveform, computed as the difference between the maxima and minima of each beat, was extracted as a measure of the amplitude of the peripheral blood volume pulse, and indirectly reflects sympathetic activity. PAT, a measure of relative changes in blood pressure and the pre-ejection period (PEP), was extracted as the time difference between R-peak of ECG to two different reference points of each PPG beat, named as PAT_{FOOT} (computed from through) and PAT_{PEAK} (computed from peak). A five-beat time constant for exponential moving averaging was used for PAT calculation.

TABLE II
ROC AUC SCORES OF CLASSIFIERS DERIVED USING LOSO-CV USING ALL FEATURES

Naïve Bayes	Random Forest	k-NN	Logistic Regression	MLP
0.52	0.76	0.56	0.80	0.71
Decision Trees	**SVM (Linear)**	**SVM (Polynomial)**	**SVM (RBF)**	**SVM (Sigmoid)**
0.62	0.80	0.66	0.65	0.98

ROC AUC = Receiver operator characteristics area under curve; k-NN = k-Nearest Neighbors; MLP = Multilayer Perceptron; SVM (Kernel) = Support Vector Machine (Kernel Function); RBF = Radial Basis Function.

PEP: PEP is a systolic time interval measured from ECG and SCG beats as an acute indicator of contractility and cardiac sympathetic activity [42]. It is computed as the time delay from ECG R-peak to the second peak in SCG signal, with a three-beat time constant for exponential moving averaging.

Respiratory measures: As the vagus nerve is heavily involved in parasympathetic control, we also extracted measures that indicate respiratory variations and are documented to dominate the parasympathetic function from the RSP waveform [43]. These measures include respiratory rate (RR), width (RW), prominence (RP). The RSP signal was detrended to avoid undesired DC offset using a sixth order polynomial for each one-minute signal segment. The rate of the inspiration peak appearance was extracted as RR (in breaths per minute). After finding the peaks, the width and prominence of each peak were also extracted as RW and RP.

E. Dataset Construction

Machine learning techniques were utilized for detecting target engagement using the continuously obtained physiological parameters. The approach consisted of determining if the physiological features could be used to correctly classify the stimulus type: active tcVNS or sham. Fig. 2b depicts the dataset construction paradigm. The physiological parameters were extracted from stimulation and post-stimulation intervals, which correspond to the data from the last minute of stimulation and one-minute data from three minutes after the stimulation ended, respectively. Then, these parameters were divided by the mean values obtained from the rest period to normalize inter-subject variability, as in our prior physiological sensing work [44]. The rest data were obtained before the stimulation protocol started, at the same body position the stimulation was carried out. The dataset for machine learning was constructed from the time-series data obtained from the normalized intervals. The extracted parameters were resampled to the length of the parameter that has the maximum length to equalize the length of each parameter within an interval for further usage of the resampled data in windowing (to obtain equal length of arrays for windowing), using an antialiasing FIR lowpass filter and compensating for the delay introduced by the filter [45]. Resampling was necessary as the length of respiratory features are approximately half of the beat-by-beat features per unit time, as human respiration and heartbeat frequency are different

(12-20 breaths per minute respiratory rate versus 70-100 beats per minute heart rate). Then, instances were created by using 10-sample sliding windows with 9-sample overlap (90% overlap). The features consisted of the mean, standard deviation (std), maximum (max), minimum (min), area under curve (auc) and slope (slp) of the extracted physiological parameters in each window. The approximate first order derivatives (differences between the adjacent elements) were also computed to generate additional parameters. The same feature calculation methodology was applied to the difference matrices (except slope). The multi-dimensional feature matrix was constructed from all extracted features and the corresponding labels as device types (tcVNS, sham). This matrix included features as columns and instances as rows, consisting of 176 features and 1780 instances (827 tcVNS, 953 sham). Overall, these features were derived from 16 time-series data (HR, PEP, PPG Amplitude, PAT_{FOOT}, PAT_{PEAK}, RR, RW, RP in two intervals), with 11 mathematical properties derived from each of them and their differences. Later, the data were standardized across each feature to have a mean of zero (subtracting the mean in the numerator in Fig. 2c) and a standard deviation of one (unit variance, by dividing by the standard deviation of the data as seen in Fig. 2c). Standardization is necessary because some parameters respond differently for the same level of stress due to the physiological nature of the measurement. (i.e., 5% decrease in PEP or 80% decrease in PPG amplitude might both indicate similar levels of stress).

F. Visualization, Feature Selection, Machine Learning

Fig. 2c summarizes the visualization, feature selection, and machine learning approach. For demonstration purposes, t-distributed Stochastic Neighbor Embedding (t-SNE) was applied to the high-dimensional standardized feature matrix to visualize the dataset in two dimensions [46, 47]. For the machine learning paradigm, our primary concerns were 1) minimizing the number of features/sensing modalities used in classification to optimize the computational power and to determine which sensing modalities are necessary for a wearable implementation (note that there are a total of 176 features from four sensors with unknown contributions to the classification task), 2) developing a realistic validation paradigm for a new subject who has just started undergoing stimulation. To address 1, we implemented a univariate feature selection paradigm based on sorting ANOVA F-statistics of features to eliminate redundant data without incurring significant loss of information [48]. Note that this feature selection step is not mandatory but optional. Assuming no need for optimization, not applying this step (simply using all features) did not cause a dramatic decrease in performance (See Table II, SVM with sigmoid kernel, which was used as our classifier). To address 2, we wanted to ensure that the training dataset are not biased for a new, incoming subject by using leave-one-subject-out cross validation (LOSO-CV). Typical methods such as K-fold cross validation would not guarantee that the training dataset does not have data from the incoming subject. The fact that there is data from the incoming subject in both the training and testing datasets will likely make the model know more about the target subject than it should. With an entirely new subject (not in the training dataset), the K-fold trained model will potentially perform poorly because it did not include data from the new subject in the training set before. LOSO-CV, on the other hand, guarantees there is no data from the incoming subject in the training dataset in model evaluation process already, hence this is a conservative approach to assess an expected performance for an incoming subject. Feature selection and LOSO-CV were implemented as follows: For each LOSO-CV loop, one subject was left out of the feature selection and classification model training, then used for testing. The procedure was repeated for each of the 26 subjects.

We determined the number of top features by implementing the feature selection and LOSO-CV and swiping the number of top features from 1 to 176. For each case, we computed the area under curve obtained from the receiver operator characteristics (ROC AUC) and plotted $k_{topfeatures}$ versus ROC AUC. Then, we determined a range for $k_{topfeatures}$ where ROC AUC is high (>0.9) and robust enough to ensure good performance and $k_{topfeatures}$ is low enough to reduce processing time.

Following the determination of feature selection and the needs for model evaluation, we explored the most effective classification architecture for predicting stimulation presence from physiological signals. Using LOSO-CV, we used the following classifiers: Naïve Bayes (for predicting baseline performance), Random Forest, k-Nearest Neighbors (k-NN), Logistic Regression, Multilayer Perceptron, Decision Trees, and Support Vector Machine (SVM) with linear, polynomial, radial bases function (RBF), and sigmoid kernels. Table II (see Results section below) lists the ROC AUC values for each classifier. We observe that SVM with a sigmoid kernel performs the best among different classifiers used, hence we decided to use this classifier.

III. Results

We performed multiple classification runs with a sigmoid SVM classifier using features from different sensors alone and their combinations. Figure S1 shows the number of features versus ROC AUC, plotted from one feature to the maximum number of features based on the sensors chosen. It is apparent that ROC AUC>0.8 could be obtained from different sensor combinations.

By using the described feature selection and machine learning methodologies, we obtained the highest ROC AUC (>0.90) and highest accuracy, precision, recall (>0.90) for four cases: By using: 1) all sensors, at least 140 top features out of 176 (Figure S1-i); 2) ECG and PPG sensors, at least 70 top features out of 88 (Figure S1-vi); 3) ECG, SCG, and PPG sensors, at least 99 top features out of 110 (Figure S1-x), and 4) ECG, PPG, and RSP sensors, at least 84 features out of 154 (Figure S1-xi). As our aim was to minimize the number of sensing modalities and features to reduce complexity for a future wearable implementation, we focused on the classification results for the second case, the combination of ECG and PPG. The features in this dataset contain HR, PPG amplitude, and PAT.

Our next goal was to analyze the results obtained from ECG and PPG. Fig. 3 details the outcomes for the features obtained from ECG and PPG. Fig. 3a shows the dimensionality reduction (t-SNE) plot grouped by tcVNS and sham clusters. A nonlinear separation exists between the device groups. Fig. 3b shows how the number of selected top features change the ROC area. A

ROC area of >0.9 can be obtained by using 70 top features out of 88 (and beyond). Fig. 3c-d summarize the machine learning outcomes for the classification using top 71 features out of 88 (note that outcomes are similar between 70 to 88). The classification resulted in 25 correctly classified subjects out of 26 total subjects (12 active tcVNS, 14 sham) with LOSO-CV. There was one false negative subject. Macro-averaged accuracy, precision, recall, and F-1 scores of 96% with 0.96 ROC AUC were obtained as the performance outcomes. Due to the feature selection method applied in each LOSO-CV loop, the features used in the classification slightly differed for each subject. Fig. 4 shows the top 5 features obtained by applying the feature selection method to the whole dataset of 88, sorted by ANOVA F-values. PPG and HR features resulted in the largest F-values. Lastly, we calculated the time complexity (the time elapsed encompassing all tasks ranging from signal processing to machine learning) for each subject as measured on a Core i7-6500U CPU @2.5GHz 12GB RAM personal laptop. We found that 2.6 (1.1) seconds (mean (SD)) were required to generate an output class for an incoming subject.

IV. Discussion

In this work, we investigated methods for detecting target engagement for tcVNS by using cardiovascular and peripheral signals from a double-blinded study. The dataset constructed from ECG and PPG signals yielded sufficient information to detect whether therapy response occurred (active device), resulting in 25 correctly classified subjects out of 26. The false negative subject (active subject classified as sham) was a female aged 24 years. The combination of other sensing modalities resulted in similar classification outcomes, slightly changing the confusion matrix. Single sensors/features do not offer the same performance as the combinations. Thus, there is not a single biomarker or pattern that could be easily recognized by the human eyes; there is likely a complex, high-dimensional relation between multiple features that requires multiple modalities.

The investigation of the most salient features reveals that features related to PPG and HR perform favorably compared to others. HR is regulated by both SNS and PNS activity [38]. Stimulation of the vagus nerve typically decreases SNS activity or increases PNS activity, also observed in the clinical portion of this study [49]. Moreover, increase in PPG amplitude was noted when tcVNS was paired with traumatic stress, compared to sham. From Fig. 4b-f, PPG features increased and mean HR decreased for the tcVNS group. Both of these changes indicate a decrease in sympathetic tone, which perhaps makes the separation between the classes possible by these features.

The analysis of machine learning performance outcomes is essential to gain an understanding regarding the translation of the methods to at-home settings: with our classifier, the accuracy, precision, recall, F-1 scores were 0.96. Thus, the classifier could provide a binary (e.g., red or green) indicator to the user following stimulation regarding whether the nerve target was successfully engaged. Among 100 stimulation administrations upon traumatic stress triggers, we would be able detect whether the stimulation happened correctly 96 times.

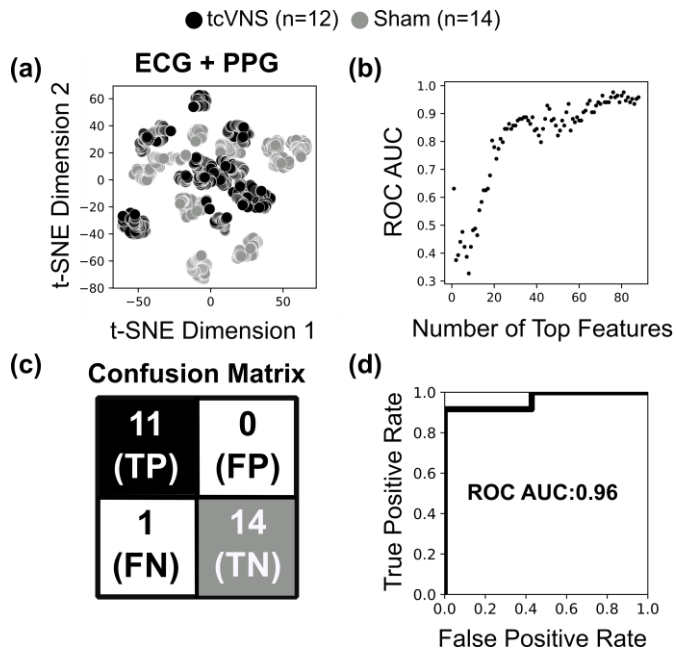

Fig. 3. Dimensionality reduction and classification outcomes for separating the stimulus types: active tcVNS and sham. (a) Dimensionality reduction applied to the high-dimensional feature matrix using t-SNE constructed from features from ECG and PPG. (b) Number of Top features selected using ANOVA F-score-based feature selection versus receiver operator characteristics (ROC) area under curve (AUC). ROC AUC is robust to Top Features from 70 to 88. c) Confusion matrix for the classifier, obtained with LOSO-CV and minimum number of features (71). (d) Receiver operator characteristics (ROC) for the classifier. A ROC area under curve (AUC) of 0.96 was obtained. Classification outcomes vary minorly with Top Features from 70 to 88.

The methods described in this study leverage physiological signals that are convenient to obtain with wearable sensing modalities, such as the smartwatches that measure ECG and PPG. The framework presented herein could remain the same and could be generalized to include other types of sensory data. Note that although the measured computation time was 2.6 seconds on average to generate an output class for an incoming subject, this was computed with a personal laptop. A final implementation of this pipeline would not require a fully embedded platform as no precision timing is necessary. The collected signals could be transferred to a Cloud-based distributed system that employs parallel processing capabilities to further reduce this 2.6-second computation time if needed.

Among the modalities we used, HR and respiratory effort information are typically collected in clinical practice, in a non-continuous manner. Therefore, we explored using only these features. Using the features obtained from HR (Figure S1-ii), RSP (Figure S1-v), and their combination (Figure S1-viii) for the classification task do not outperform the reported results.

Other measures that might indicate stimulation presence from the literature are the use of microneurography, pupil size, serum cytokines (anti-inflammatory effects as quantified by tumor necrosis alpha, interleukin-6), heart rate variability (HRV), or evoked potentials resulting from stimulation [10, 20, 50, 51]. Microneurography and serum cytokines are not non-invasive, not pursuant to the goal of real-time identification, and are confined to clinical settings, electroencephalogram [52] requires bulky multi-channel equipment, while pupil size was not found to give favorable results for stimulation in prior work.

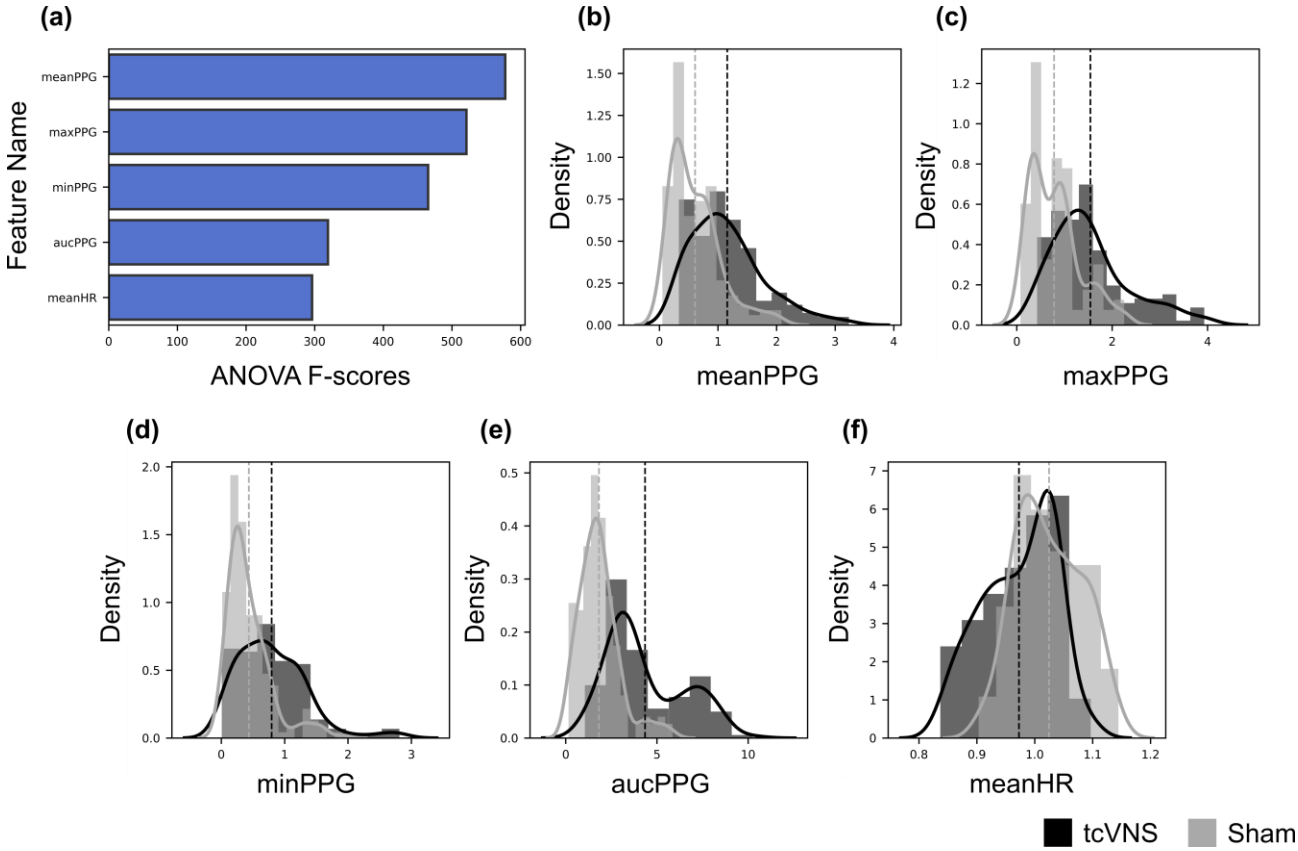

Fig. 4. Top 5 features sorted by ANOVA F-values, their histograms, and kernel density estimates grouped by classes. The top features were calculated from the full feature set obtained from ECG and PPG sensors. The dashed lines indicate the mean of the class.

As for HRV, this measure does not have a continuous nature and requires a long-term clean recording to compute a single number based on the variability in R-R intervals. An auricular stimulation study noted changes in HRV with long-term recordings [10], while some studies did not note changes in HRV [17, 53].

A number of limitations exist for the study described here. First, this study employs a reactive approach. Traumatic stress requires a flashback occurrence: the subjects should remember their memories to be stressed, which requires some time starting from the introduction of the traumatic recording. We applied stimulation immediately after the traumatic stress recording as our prior brain imaging studies demonstrated that the arousal persists following rumination of the traumatic script recording [54, 55]. Future studies should consider sweeping the timing of the stimulation, which might possibly downregulate the autonomic reactivity even before reaching a peak stressed state, based on prior preclinical studies [56, 57]. Second, this study did not have continuous blood pressure recordings, hence we could not use blood pressure-related measures except pulse arrival time (PAT) obtained from ECG and PPG, which is a measure related to *both* continuous blood pressure and cardiac contractility [58]. Nevertheless, analyses on the effects of auricular or cervical stimulation on cardiovascular and autonomic function have produced mixed outcomes throughout many studies that use basic vital signals such as HR, HRV, or BP [10, 17, 18, 21, 23, 53], hence these basic measures are not likely to be sufficient for monitoring stimulation presence. Lastly, in the current study, we obtained a transmissive PPG signal that requires the photodiode (PD) and light emitting diode (LED) combination to be at the opposite sides of the skin. A reflective PPG sensor (both PD and LED on the same side of the skin) might be more appropriate for a comfortable and minimally obtrusive wearable device.

V. Conclusion

This study demonstrates the first effort to provide real-time inputs that could be used in tcVNS therapy response, and further for closed-loop modulation for at-home tcVNS technologies. Multimodal signal fusion might be a viable approach in determining whether the stimulation occurred as expected. In addition to the investigation of individual parameters, sensor fusion could be instrumental in translating tcVNS to unsupervised settings to improve therapy response in a home-based setting. ECG and PPG sensing appear to provide relevant information regarding the stimulation delivery, and both signals could be obtained noninvasively, and are also prevalent in commercially available wearable sensing devices. The methods presented herein could thus be deployed in wearables allowing for a convenient home-based approach to supporting accurate and effective delivery of noninvasive vagal nerve stimulation therapies.

Future studies should focus on the following to facilitate successful translation to clinical practice. Continuous blood pressure should be measured as an additional physiological parameter to be studied as it includes important vascular information; effects of the time of stimulation (before or during traumatic stimuli) on efficacy should be determined; and the ability of indices derived from reflective PPG sensors to capture

the key physiological information should be quantified. Additionally, stimulation induces parameter-specific changes in physiology [11]. Therefore, regression models of stimulation parameters onto the changes in cardiovascular and peripheral measures could provide a better understanding for inter-patient variability in noninvasive VNS studies.

REFERENCES

[1] H. R. Berthoud and W. L. Neuhuber, "Functional and chemical anatomy of the afferent vagal system," *Auton Neurosci,* vol. 85, no. 1-3, pp. 1-17, Dec 20 2000.

[2] S. Breit, "Vagus nerve as modulator of the brain–gut axis in psychiatric and inflammatory disorders," *Frontiers in psychiatry,* vol. 9, no. 44, 2018.

[3] R. L. Johnson and C. G. Wilson, "A review of vagus nerve stimulation as a therapeutic intervention," *Journal of inflammation research,* vol. 11, pp. 203–213, 2018.

[4] E. Frangos and B. R. Komisaruk, "Access to Vagal Projections via Cutaneous Electrical Stimulation of the Neck: fMRI Evidence in Healthy Humans," *Brain Stimul,* vol. 10, no. 1, pp. 19-27, Jan - Feb 2017.

[5] N. Yakunina, S. S. Kim, and E. C. Nam, "Optimization of Transcutaneous Vagus Nerve Stimulation Using Functional MRI," *Neuromodulation,* vol. 20, no. 3, pp. 290-300, Apr 2017.

[6] Z. D. Deng, S. M. McClintock, N. E. Oey, B. Luber, and S. H. Lisanby, "Neuromodulation for mood and memory: from the engineering bench to the patient bedside," *Curr Opin Neurobiol,* vol. 30, pp. 38-43, Feb 2015.

[7] W. P. Teo *et al.,* "Does a Combination of Virtual Reality, Neuromodulation and Neuroimaging Provide a Comprehensive Platform for Neurorehabilitation? - A Narrative Review of the Literature," *Front Hum Neurosci,* vol. 10, p. 284, 2016.

[8] J. Redgrave *et al.,* "Safety and tolerability of Transcutaneous Vagus Nerve stimulation in humans; a systematic review," *Brain Stimul,* vol. 11, no. 6, pp. 1225-1238, Nov - Dec 2018.

[9] E. Ben-Menachem, D. Revesz, B. J. Simon, and S. Silberstein, "Surgically implanted and non-invasive vagus nerve stimulation: a review of efficacy, safety and tolerability," *Eur J Neurol,* vol. 22, no. 9, pp. 1260-8, Sep 2015.

[10] J. A. Clancy, D. A. Mary, K. K. Witte, J. P. Greenwood, S. A. Deuchars, and J. Deuchars, "Non-invasive vagus nerve stimulation in healthy humans reduces sympathetic nerve activity," *Brain Stimul,* vol. 7, no. 6, pp. 871-7, Nov-Dec 2014.

[11] B. W. Badran *et al.,* "Short trains of transcutaneous auricular vagus nerve stimulation (taVNS) have parameter-specific effects on heart rate," *Brain Stimul,* vol. 11, no. 4, pp. 699-708, Jul - Aug 2018.

[12] R. G. Garcia *et al.,* "Modulation of brainstem activity and connectivity by respiratory-gated auricular vagal afferent nerve stimulation in migraine patients," *Pain,* vol. 158, no. 8, pp. 1461-1472, Aug 2017.

[13] H. M. Hamer and S. Bauer, "Lessons learned from transcutaneous vagus nerve stimulation (tVNS)," *Epilepsy Res,* vol. 153, pp. 83-84, Jul 2019.

[14] J. Fang *et al.,* "Transcutaneous Vagus Nerve Stimulation Modulates Default Mode Network in Major Depressive Disorder," *Biol Psychiatry,* vol. 79, no. 4, pp. 266-73, Feb 15 2016.

[15] P. Rong *et al.,* "Effect of transcutaneous auricular vagus nerve stimulation on major depressive disorder: A nonrandomized controlled pilot study," *J Affect Disord,* vol. 195, pp. 172-9, May 2016.

[16] D. G. Lamb, E. C. Porges, G. F. Lewis, and J. B. Williamson, "Non-invasive Vagal Nerve Stimulation Effects on Hyperarousal and Autonomic State in Patients with Posttraumatic Stress Disorder and History of Mild Traumatic Brain Injury: Preliminary Evidence," *Front Med (Lausanne),* vol. 4, p. 124, 2017.

[17] A. M. Burger *et al.,* "Mixed evidence for the potential of non-invasive transcutaneous vagal nerve stimulation to improve the extinction and retention of fear," *Behav Res Ther,* vol. 97, pp. 64-74, Oct 2017.

[18] A. M. Burger, B. Verkuil, I. Van Diest, W. Van der Does, J. F. Thayer, and J. F. Brosschot, "The effects of transcutaneous vagus nerve stimulation on conditioned fear extinction in humans," *Neurobiol Learn Mem,* vol. 132, pp. 49-56, Jul 2016.

[19] A. M. Burger, W. Van der Does, J. F. Thayer, Brosschot;, J. F.;, and B. Verkuil, "Transcutaneous Vagus Nerve Stimulation Reduces Spontaneous but not Induced Negative Thought Intrusions in High Worriers," *Biological Psychology,* 2019.

[20] I. Lerman *et al.,* "Noninvasive Transcutaneous Vagus Nerve Stimulation Decreases Whole Blood Culture-Derived Cytokines and Chemokines: A Randomized, Blinded, Healthy Control Pilot Trial," *Neuromodulation,* vol. 19, no. 3, pp. 283-90, Apr 2016.

[21] C. Brock *et al.,* "Transcutaneous cervical vagal nerve stimulation modulates cardiac vagal tone and tumor necrosis factor-alpha," *Neurogastroenterol Motil,* vol. 29, no. 5, May 2017.

[22] M. L. Oshinsky, A. L. Murphy, H. Hekierski, Jr., M. Cooper, and B. J. Simon, "Noninvasive vagus nerve stimulation as treatment for trigeminal allodynia," *Pain,* vol. 155, no. 5, pp. 1037-42, May 2014.

[23] S. P. Chen *et al.,* "Vagus nerve stimulation inhibits cortical spreading depression," *Pain,* vol. 157, no. 4, pp. 797-805, Apr 2016.

[24] S. J. Hellawell and C. R. Brewin, "A comparison of flashbacks and ordinary autobiographical memories of trauma: Content and language," *Behaviour Research and Therapy,* vol. 42, no. 1, pp. 1-12, 2004.

[25] C. R. Brewin, "Re-experiencing traumatic events in PTSD: New avenues in research on intrusive memories and flashbacks," *European Journal of Psychotraumatology,* vol. 6, no. 1, p. 27180, 2015.

[26] N. A. Kline and J. L. Rausch, "Olfactory precipitants of flashbacks in posttraumatic stress disorder," *The Journal of clinical psychiatry,* 1985.

[27] O. Van der Hart and B. Friedman, "Trauma, dissociation and triggers: Their role in treatment and emergency psychiatry," *Emergency psychiatry today,* pp. 137-142, 1992.

[28] B. M. Cortese, K. Leslie, and T. W. Uhde, "Differential odor sensitivity in PTSD: Implications for treatment and future research," (in eng), *Journal of affective disorders,* vol. 179, pp. 23-30, 2015.

[29] J. Spence *et al.,* "Characteristics and Treatment Preferences of People with Symptoms of Posttraumatic Stress Disorder: An Internet Survey," *PLOS ONE,* vol. 6, no. 7, p. e21864, 2011.

[30] N. Z. Gurel *et al.,* "Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress," *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation,* vol. 13, no. 1, pp. 47-59, 2020.

[31] M. D. Johnson *et al.,* "Neuromodulation for brain disorders: challenges and opportunities," *IEEE Trans Biomed Eng,* vol. 60, no. 3, pp. 610-24, Mar 2013.

[32] N. Z. Gurel *et al.,* "Toward Closed-Loop Transcutaneous Vagus Nerve Stimulation using Peripheral Cardiovascular Physiological Biomarkers: A Proof-of-Concept Study," presented at the IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Las Vegas, NV, 2018.

[33] N. Z. Gurel *et al.,* "Timing Considerations for Noninvasive Vagal Nerve Stimulation in Clinical Studies," presented at the American Medical Informatics Association Annual Symposium, Washington, DC, 2019.

[34] A. P. Association, *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.

[35] M. B. First and M. Gibbon, "The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II)," in *Comprehensive handbook of psychological assessment,* vol. 2, M. J. H. D. L. Segal, Ed. Hoboken, NJ, US: John Wiley & Sons Inc., 2004, pp. 134-143.

[36] L. Sornmo, *Bioelectrical Signal Processing in Cardiac and Neurological Applications,* 1st ed. Burlington, MA: Elsevier, 2005.

[37] N. Z. Gurel *et al.,* "Comparison of autonomic stress reactivity in young healthy versus aging subjects with heart disease," *PLOS ONE,* vol. 14, no. 5, p. e0216278, 2019.

[38] M. N. Levy, "Sympathetic-parasympathetic interactions in the heart," *Circ Res,* vol. 29, no. 5, pp. 437-45, Nov 1971.

[39] J. L. Ardell, P. S. Rajendran, H. A. Nier, B. H. KenKnight, and J. A. Armour, "Central-peripheral neural network interactions evoked by vagus nerve stimulation: functional consequences on control of cardiac function," *Am J Physiol Heart Circ Physiol,* vol. 309, no. 10, pp. H1740-52, Nov 15 2015.

[40] K. Yamakawa *et al.*, "Vagal nerve stimulation activates vagal afferent fibers that reduce cardiac efferent parasympathetic effects," *Am J Physiol Heart Circ Physiol*, vol. 309, no. 9, pp. H1579-90, Nov 2015.

[41] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol Meas*, vol. 28, no. 3, pp. R1-39, Mar 2007.

[42] O. T. Inan *et al.*, "Ballistocardiography and seismocardiography: a review of recent advances," *IEEE J Biomed Health Inform*, vol. 19, no. 4, pp. 1414-27, Jul 2015.

[43] G. G. Berntson *et al.*, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623-48, Nov 1997.

[44] N. Z. Gurel, H. Jung, S. Hersek, and O. T. Inan, "Fusing Near-Infrared Spectroscopy With Wearable Hemodynamic Measurements Improves Classification of Mental Stress," *IEEE Sensors Journal*, vol. 19, no. 19, pp. 8522-8531, 2019.

[45] M. MathWorks, "6.1 Signal Processing Toolbox Documentation," *MathWorks Inc., Natick, MA, USA*, 2001.

[46] Laurens van der Maaten and G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, vol. 9, pp. 2579-2605, 2008.

[47] F. Pedregosa, "Scikit-learn: Machine learning in Python," *Journal of machine learning research*, vol. 12, pp. 2825-2830, 2011.

[48] Yoav Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the royal statistical society. Series B (Methodological)*, vol. 57, no. 1, pp. 289-300, 1995.

[49] N. Z. Gurel *et al.*, "Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress," *Brain Stimulation*, 2019/08/06/ 2019.

[50] M. Keute, M. Demirezen, A. Graf, N. G. Mueller, and T. Zaehle, "No modulation of pupil size and event-related pupil response by transcutaneous auricular vagus nerve stimulation (taVNS)," *Scientific Reports*, vol. 9, no. 1, p. 11452, 2019/08/07 2019.

[51] C. Mugruza-Vassallo and D. Potter, "Context Dependence Signature, Stimulus Properties and Stimulus Probability as Predictors of ERP Amplitude Variability," (in English), *Frontiers in Human Neuroscience*, Original Research vol. 13, no. 39, 2019-February-26 2019.

[52] C. T. Skau *et al.*, "Retraction Notice to: FMN2 Makes Perinuclear Actin to Protect Nuclei during Confined Migration and Promote Metastasis," *Cell*, vol. 173, no. 2, p. 529, Apr 5 2018.

[53] Verkuil B and B. AM, "Transcutaneous vagus nerve stimulation does not affect attention to fearful faces in high worriers," *Behaviour research and therapy*, vol. 1, no. 113, pp. 25-31, 2019.

[54] J. D. Bremner, M. Narayan, L. H. Staib, S. M. Southwick, T. McGlashan, and D. S. Charney, "Neural correlates of memories of childhood sexual abuse in women with and without posttraumatic stress disorder," *American Journal of Psychiatry*, vol. 156, no. 11, pp. 1787-1795, 1999.

[55] J. D. Bremner *et al.*, "Positron emission tomographic imaging of neural correlates of a fear acquisition and extinction paradigm in women with childhood sexual-abuse-related post-traumatic stress disorder," *Psychol Med*, vol. 35, no. 6, pp. 791-806, Jun 2005.

[56] D. F. Pena, J. E. Childs, S. Willett, A. Vital, C. K. McIntyre, and S. Kroener, "Vagus nerve stimulation enhances extinction of conditioned fear and modulates plasticity in the pathway from the ventromedial prefrontal cortex to the amygdala," *Front Behav Neurosci*, vol. 8, p. 327, 2014.

[57] N. D. Engineer, J. R. Riley, J. D. Seale, W. A. Vrana, and J. A. Shetake, "Reversing pathological neural activity using targeted plasticity," *Nature*, vol. 470, no. 7332, p. 101, 2011.

[58] R. Mukkamala *et al.*, "Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice," *IEEE Trans Biomed Eng*, vol. 62, no. 8, pp. 1879-901, Aug 2015.