Sensitive Physiological Indices of Pain based on Differential Characteristics of Electrodermal Activity

Youngsun Kong, Member, IEEE, Hugo F. Posada-Quintero, Member, IEEE, and Ki H. Chon, Member, IEEE

***Abstract*— *Objective:* Electrodermal activity (EDA) has been widely used to assess human response to stressful stimuli, including pain. Recently, spectral analysis of EDA has been found to be more sensitive and reproducible for assessment of sympathetic arousal than traditional indices (e.g., tonic and phasic components). However, none of the aforementioned analyses incorporate the differential characteristics of EDA, which could be more sensitive to capturing fast-changing dynamics associated with pain responses. *Methods:* We have tested the feasibility of using the derivative of phasic EDA and the modified time-varying spectral analysis of EDA. Sixteen subjects underwent four levels of pain stimulation using electric stimulation. Five-second segments of EDA were used for each level of stimulation, and pre-stimulation segments were considered stimulation level 0. We used support vector machines with the radial basis function kernel and multi-layer perceptron for three different scenarios of stimulation-level classification tasks: five stimulation levels (four levels of stimulation plus no stimulation); low, medium, and high pain stimulation (stimulation levels 0-1, 2, and 3-4, respectively); and high stimulation levels (stimulation levels 3-4) vs. no stimulation. *Results:* The maximum balanced accuracies were 44% (five stimulation levels), 63% (for low, medium, and high pain stimulation), and 87% (sensitivity 83% and specificity 89%, for high stimulation vs. no stimulation). *Conclusion:* The differential characteristics of EDA contributed highly to the accuracy of pain stimulation level detection of the classifiers. The external validity dataset was not considered in the study. *Significance:* Our approach has the potential for accurate pain quantification using EDA.**

***Index Terms*—electrodermal activity, machine learning, pain, phasic component, spectral analysis, tonic component**

I. Introduction

Intensity of pain is commonly assessed using scales in which the person suffering pain reports a number, a color code, a facial expression, that best describes the level of pain being experienced [1], [2]. Although these scales are widely used in practice [3], they are highly variable and subjective. This subjectivity can be exploited by patients, in the worst case, to deceive medical providers to obtain access to painkillers such as opioids, which can potentially cause drug misuse and overdose [4], [5]. If pain can be objectively measured using physiological signals, this risk can be minimized. Furthermore, as the scales are trust-based, if the doctor does not trust the patient, or has a racial or gender bias, he or she could underestimate a patient's pain sensations, affecting the treatment. An objective measure of pain would guide doctors in providing appropriate pain management and treatment.

Human response to pain includes a set of autonomic activations [6]. Specifically, the sympathetic nervous system reaction is the most sensitive to pain stimulus intensity [7], [8]. Electrodermal activity (EDA) is a sensitive measure of sympathetic function [9]–[11], which makes it an ideal sensor for pain detection. Several features in both the time- and frequency-domain of EDA signals have been proposed to evaluate the sympathetic arousal [9]. However, questions remain of whether the features can successfully detect pain responses in EDA given the short-lasting and instantaneous characteristics of some pain. We hypothesize that differential characteristics of EDA are more suitable for quantifying pain stimuli in EDA signals, as they extract the faster-changing dynamics of the signal.

The EDA signal has two salient components, the tonic (slow, low frequency) and phasic (rapid, higher frequency) [9], [12], which have been shown to be concomitantly sensitive to sympathetic arousal [13]. There are several algorithms available for tonic/phasic decomposition of EDA [9]. These algorithms also estimate the phasic drivers that represent the underlying sympathetic activation which are represented in the dynamics of EDA. Their techniques have been tested using various datasets but mostly related to emotion and stress. It is not clear if they are also sensitive to various pain stimulation levels.

There is no clear understanding of the mechanisms involved in the interplay between pain and the autonomic reactions, but it is well known that noxious painful stimuli cause autonomic reactions, and that autonomic activity affects pain responses [14], [15]. Furthermore, the intensity of the pain stimulus [16]–[19], rather than the pain experienced by the subject, has

This work was funded by NIH R21 DE029563-01.

All authors are with the Biomedical Engineering Department, University of Connecticut, Storrs, CT 06268 USA

(e-mail: youngsun.kong@uconn.edu, hugo.posada-quintero@uconn.edu, ki.chon@uconn.edu).

stronger correlation to autonomic reactions [7]. In the rank of importance of autonomic reactions to pain, the sympathetic function is the most sensitive to stimulus intensity [7], [8].

EDA is the measure of the response due to modulated skin conductance produced by sweat gland activity (expressed in microsiemens [μS]) [12], [20]–[22]. It is frequently measured in plantar and palmar sites because sweat glands in those sites are denser than other body sites [22], [23]. Several studies showed that EDA is a good measurement tool for evaluating sympathetic arousal [24], [25]. Both tonic and phasic (which comprises the skin conductance responses, SCRs) components are widely used to assess sympathetic arousal [12], [26]–[29]. Tonic/phasic decomposition has been studied by various research groups. Among the available algorithms developed for such a task, continuous decomposition analysis [30], discrete decomposition analysis [30], dynamic causal modeling [31], the convex optimization approach [32], and the sparse deconvolution approach [33] are the most widely used. Furthermore, several EDA analysis methods for sympathetic arousal tracking have recently been developed and their sensitivity and reproducibility have been demonstrated to be outstanding, including a sparse decomposition approach with physiological priors [34], a marked point process filtering approach [35], a mixed filter algorithm [36], an index of sympathetic control based on the power spectral analysis of EDA [37], and variable frequency complex demodulation (termed TVSymp) [38]. We selected TVSymp for analysis because of its highly sensitive and consistency for sympathetic tone assessment (especially pain) [27], [38]–[40].

In this study, we aimed to test features based on the differential characteristics of EDA, along with the aforementioned approaches for EDA decomposition and spectral analysis, for classification and regression of various stimulation levels.

II. Materials and Methods

A. Subjects

The study protocol was approved by the Institutional Review Board of the University of Connecticut. Sixteen healthy volunteers (7 males and 9 females) in the age range 25.6 ± 4.8 (mean ± standard deviation), were enrolled in this study.

B. Stimulation

Pain stimulation was induced using electric pulses. A programmable stimulus isolation adapter (STMISOC, BIOPAC Systems, Inc.) was used to adjust stimulation pulses. The stimulation was delivered through two disposable silver/silver chloride (Ag/AgCl) electrodes located on the right forearm (5 cm apart), following the recommendations from the safety advisory board [41]. For each subject, we first determined the maximum stimulation level that evoked a pain level of 7 out of 10 visual analogue scales, labelled as Smax. The stimulation pulse amplitude was variable, while the pulse width was fixed to 10 *ms*. We defined 4 stimulation intensities by varying the pulse amplitude to 25% ("level 1"), 50% ("level 2"), 75% ("level 3") and 100% ("level 4") of Smax. A pseudo-random sequence of 40 stimuli with the four different amplitudes was programmed. The inter-stimulus interval was randomly set between 10-30 seconds. After each stimulus, subjects were asked to report their pain level (0 to 10). All of the 5-second windows before each stimulus were used as "level 0" instances (no stimulation).

C. Devices

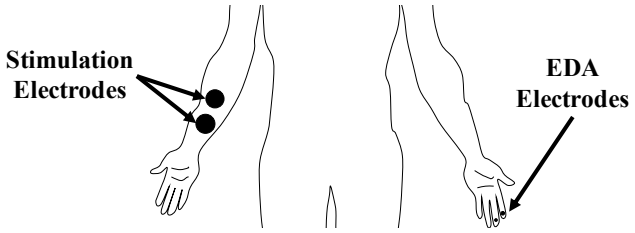


Fig. 1. Placement of stimulation and EDA electrodes.

A galvanic skin response device (GSR MP160) and an amplifier (BIONAMADIX 2CH Amp) were used to collect EDA data from the subject's left middle and index fingers during the stimulation. The skin where electrodes were placed was cleaned with alcohol before the experiment. The EDA signals were recorded at 500 Hz. Figure 1 illustrates the placement of stimulation and EDA electrodes.

D. Signal processing

1) Decomposition of EDA into tonic and phasic components

In the pre-processing stage, EDA data were down-sampled to 4 Hz, then a median filter (1-sec window) and a low-pass FIR filter (cut-off frequency = 1Hz) were applied. The traditional processing of EDA signals consists of decomposition into tonic and phasic components. We used five different techniques for tonic/phasic decomposition of EDA: continuous decomposition analysis (CDA) [30], discrete decomposition analysis (DDA) [30], dynamic causal modeling (DCM) [31], the convex optimization approach (cvxEDA) [32], and the sparse deconvolution approach (sparsEDA) [33]. MATLAB implementations of these algorithms are available online [9]. We used the matching pursuit algorithm implementation of DCM. We also tried the VBA implementation of DCM, but those results were not included as we obtained almost identical results to those obtained with the matching pursuit algorithm [42]. Because DCM does not provide the tonic component, we simply subtracted estimated phasic components from EDA signals to obtain the tonic component. The features for stimulation-level classification were the mean levels for the phasic components and the maximum levels for the phasic driver components of EDA obtained with each approach over the 5-sec windows starting right after each stimulus.

2) Derivative of phasic EDA

We have also computed the derivative of the phasic component of EDA as an index of pain. We hypothesize that in contrast to other types of stimulus (e.g. stress) the phasic component of EDA changes faster in response to pain. We

aimed to test if the derivative of phasic EDA (dphEDA) will improve the performance of stimulation-level classification. We computed dphEDA for each of the five approaches using the 5-point stencil central finite differences equation [43], for each sample n, as follows:

$$dphEDA(n) = \frac{P(n-2)-8\cdot P(n-1)+8\cdot P(n+1)-P(n+2)}{12\cdot(1/fs)} \quad (1)$$

where fs is the sampling frequency (4 Hz), and P is the phasic component of EDA obtained with the five different decomposition approaches. We have chosen equation (1), which uses a moving window of 1.25 s, by considering the sampling rate and the typical rising time of SCRs (between 1 and 5 seconds) [9].

3) The modified time-varying index of EDA

In addition, we obtained an index using the spectral analysis of EDA to quantify the elicited sympathetic response to pain using the approach reported in [38], modified by including a differential computation [40]. For this analysis, EDA signals were down-sampled to 2 Hz and then high-pass filtered at 0.01 Hz, as the frequency ranges of EDA response elicited by the sympathetic functions are between 0.045-0.25 Hz [37]. Variable frequency complex demodulation (VFCDM) [44], [45] was used to compute the time-frequency spectra (TFS) of EDA signals, and is briefly described here. The first step is the complex demodulation (CDM) to obtain an estimate of the TFS. A bank of low pass filters decomposes the signal into a set of signals with spectral content limited to certain bands. The Hilbert transform is used to convert the signals into analytic signals. Finally, the instantaneous amplitude, frequency, and phase is obtained for each signal. For the mathematical function CDM, let $x(t)$ to be a narrow-band sinusoidal signal oscillation whose spectrum is centered at f_0, with instantaneous amplitude $A(t)$, phase $\varphi(t)$, and the direct current component $dc(t)$ defined as:

$$x(t) = dc(t) + A(t)cos(2\pi f_0 t + \varphi(t)). \quad (2)$$

For a given center frequency, we can extract the instantaneous amplitude information $A(t)$ and phase information $\varphi(t)$ by multiplying $x(t)$ by $e^{-j2\pi f_0 t}$, which results in the following:

$$z(t) = x(t)e^{-j2\pi f_0 t}$$
$$= dc(t)e^{-j2\pi f_0 t} + \left(\frac{A(t)}{2}\right)e^{j\varphi(t)} + \left(\frac{A(t)}{2}\right)e^{-j(4\pi f_0 t+\varphi(t))}. \quad (3)$$

A leftward shift by $e^{-j2\pi f_0 t}$ moves the center frequency f_0 to zero frequency in the spectrum of $z(t)$. If $z(t)$ is subjected to an ideal low-pass filter with a cutoff frequency $f_c < f_0$, then the filtered signal $z_{1P}(t)$ will contain only the component of interest, and we obtain:

$$z_{1P}(t) = \left(\frac{A(t)}{2}\right)e^{j\varphi(t)} \quad (4)$$
$$A(t) = 2|z_{1P}(t)| \quad (5)$$
$$\varphi(t) = tan^{-1}\left(\frac{Im(z_{1P}(t))}{Re(z_{1P}(t))}\right) \quad (6)$$

Consider the case of a non-fixed modulating frequency which varies as a function of time. For that case, the signal $x(t)$ can be written in the following form:

$$x(t) = dc(t) + A(t)\left(\int_0^t cos(2\pi f(\tau)d\tau + \varphi(t))\right). \quad (7)$$

In the same way as in (2) and (3), multiplying (7) by $e^{-j\int_0^t 2\pi f(\tau)d\tau}$ produces instantaneous amplitude, $A(t)$, and instantaneous phase, $\varphi(t)$ [46], so that:

$$z(t) = x(t)e^{-j\int_0^t 2\pi f(\tau)d\tau}$$
$$= dc(t)e^{-j\int_0^t 2\pi f(\tau)d\tau} + \frac{A(t)}{2}e^{j\varphi(t)} + \frac{A(t)}{2}e^{-j\int_0^t 4\pi f(\tau)d\tau} \quad (8)$$

From (8), when $z(t)$ is filtered with an ideal low-pass filter with a cutoff frequency $f_c < f_0$, then the filtered signal $z_{1P}(t)$ will be obtained with the same instantaneous amplitude A(t) and phase $\varphi(t)$ as provided in (5) and (6). The instantaneous frequency, as reported previously [47], is given by:

$$f(t) = f_0 + \frac{1}{2\pi}\frac{d\varphi(t)}{dt} \quad (9)$$

We have used the EDA signal as $x(t)$. The spectral content of the resulting components depends on the sampling frequency (fs). Given that fs = 2 Hz, the spectral frequencies are centered on 0.04 Hz, 0.12 Hz, 0.2 Hz, 0.28 Hz, 0.36 Hz, 0.44 Hz, 0.52 Hz, 0.6 Hz, 0.68 Hz, 0.76 Hz, 0.84 Hz and 0.92 Hz. We used the components whose spectral content lies in the frequency range that was found to correspond to the sympathetic dynamics (0.045-0.25 Hz) [37]. For this reason, time-varying spectral amplitudes in the second and third components were summed together to obtain an estimated reconstructed EDA signal ($X'(t)$), which was then normalized to unit variance (divided by its standard deviation). Its instantaneous amplitude was computed using the Hilbert transform [48], as follows:

$$Y'(t) = \frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{X\prime(\tau)}{t-\tau}d\tau \quad (10)$$

where P indicates the Cauchy principal value. $X'(t)$ and $Y'(t)$ form the complex conjugate pair, so we can define an analytic signal, $Z(t)$, as

$$Z(t) = X'(t) + iY'(t) = a(t)e^{j\theta(t)} \quad (11)$$

in which

$$a(t) = [X'^2(t) + Y'^2(t)]^{1/2}$$
$$\theta(t) = arctan(Y'(t)/X'(t)) \quad (12)$$

The resulting $a(t)$ is the instantaneous amplitude of $Z(t)$. For computing the modified time-varying index of EDA, MTVSymp, the difference in value of $a(t)$ with respect to the mean value of $a(t)$ for the previous 5-sec window was computed. This process enhances the changes in the signal and partially removes other possible sources of arousal (e.g., underlying stress). Negative values were set to 0.

$$MTVSymp = max\left(0, a(t) - mean\left(a(t-w:t)\right)\right) \quad (13)$$

where w is a time window. Several values of w were tested and w was finally set to 5 seconds. The resulting MTVSymp for a given subject is shown in Figure 3 (b). This value constitutes our intended index of sympathetic reaction to pain. Notice that EDA was normalized by its standard deviation in the process of computing MTVSymp, making this parameter dimensionless.

E. Statistical analysis between stimulation levels

We evaluated the differences between pain stimulation levels (0 through 4), for the phasic driver, the phasic and tonic components obtained using the five decomposition approaches, plus the differential characteristics dphEDA and MTVSymp. We tested the normality of the measures of EDA for the different stimulation levels using the Kolmogorov-Smirnov test [49]–[51]. As all data were non-normally distributed, we used Dunn's test to compare parameters at different stimulation levels. Dunn's test is a non-parametric analog-to-multiple-pairwise t-test following rejection of an ANOVA null hypothesis [52]. A p-value < 0.05 was considered statistically significant.

F. Stimulation level classification and regression

As stimulus levels of pain are ordinal, ordinal regression (or classification) should be considered, however, typical classification approaches will perform similarly because the features are correlated with stimulus levels of pain. We have five groups, as we have stimulation levels from 0 to 4. For stimulation-level classification, a support vector machine with radial basis function kernel (R-SVM) and multi-layer perceptron (MLP) method were tested. For R-SVM, parameter C and gamma were selected from 10^k (k varied from -2 to 3 and from -4 to 1, respectively), based on the balanced accuracy for each fold using grid-search cross-validation with five-subject fold. The one-vs.-rest approach was applied. For MLP, four parameters were optimized based on cross-entropy loss for each fold using grid-search cross-validation with five-subject fold. The number of hidden layers were chosen between 1-4 with 100 hidden units. The activation functions were chosen among logistic, tanh, and rectifier unit. The optimizers were selected among the stochastic gradient decent (SGD), Adam, and limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm [53]–[55]. The initial learning rates were tested among 0.0001, 0.001, and 0.01. The learning rate was divided by 5 when two consecutive epochs failed to decrease for the SGD optimizer. The maximum number of epochs was set as 200. These algorithms were used to classify pain stimulation levels using phasic components only, including dphEDA, including MTVSymp, and including both. Given the reduced size of the dataset, we used leave-one-subject-out cross-validation to evaluate the performance of the ML models [56]. In addition to classification, we also tested R-SVM and MLP regressors. Their parameters were selected using the same method as for the classification approach, based on the coefficient of determination. To compensate for imbalanced datasets, we up-sampled the dataset using SVM-SMOTE [57], [58]. Moreover, baseline classifiers and regressors (Zero-R) were trained to observe the baseline performance as a benchmark for other classifiers and regressors. The Zero-R classifiers and regressors predicted stimulus levels to be the most frequent label in the training set and the mean of the training set, respectively. Scikit-learn and imbalanced-learn Python packages were used for both classification and regression, and the oversampling technique, respectively [59], [60].

G. Classification and regression performance evaluation

The dataset is imbalanced, as every sample of a given stimulation level has a corresponding level 0 sample. In other words, level 0 samples comprise 50% of the dataset. For this reason, instead of classification accuracy, we have computed mean balanced accuracy as the overall measure of performance, calculated using the number of correct classifications into the given groups (stimulation level 0 to 4), divided by the total number of classifications performed (number of subjects). In addition to the classification metric, we also calculated regression evaluation metric macro-averaged root-mean-square-error (M-RMSE) and the mean absolute error (M-MAE), as the target (i.e., stimulus levels) to be classified is ordinal [61], [62]. M-RMSE and M-MAE were calculated by averaging RMSE and MAE of each class, in order to compensate for the imbalanced dataset. For evaluating performance of regressors, the coefficient of determination, M-RMSE, and M-MAE were calculated.

III. RESULTS

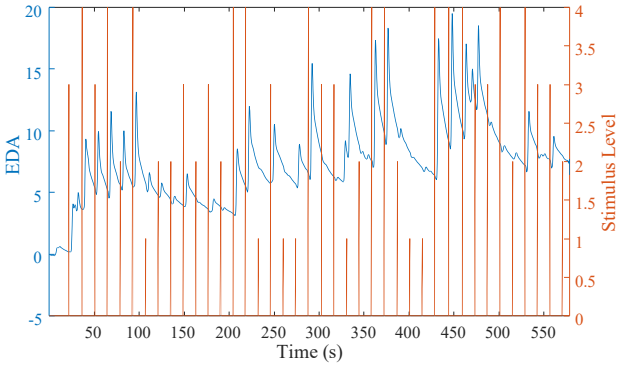


Fig. 2. EDA and stimulation level.

Figure 2 includes the collected raw EDA and stimulation levels for a given subject. The stimulation level (red) has four different levels, shown in the right y-axis of the figure. The EDA response ranged from 0 to about 20 microsiemens (μS). Noticeably, most of the stimulation pulses elicited one or more SCRs. SCRs of different amplitudes can be observed throughout the stimulation. Figure 3 shows a segment of the resulting dphEDA of cvxEDA and MTVSymp for a given subject. Steep rise on the phasic component of EDA, dphEDA, and MTVSymp after each stimulus can be observed.

Figure 4 shows a segment of the resulting tonic and phasic components of EDA using CDA, DDA, DCM, cvxEDA, and sparsEDA. The tonic component estimations by cvxEDA and

TABLE I
COMPARISON OF EDA DECOMPOSITION METHODS WITH MEAN VALUES OF EACH FEATURE (MEAN ± STANDARD DEVIATION)

Approach	*Component*	*Level 0*	*Level 1*	*Level 2*	*Level 3*	*Level 4*
CDA	phasic (μS)	0.65±0.61	0.52±0.58 [0]	0.57±0.47	0.79±0.66 [1,2]	1.1±1.2 [0,1,2]
	tonic (μS)	7.4±5	6.8±5.1	7.2±5.1	8±5.3	7.6±4.6
	dphEDA (μS)	-0.07±0.13	0.066±0.13 [0]	0.11±0.13 [0]	0.18±0.19 [0,1]	0.31±0.32 [0,1,2]
	phasic driver (a.u.)	0.44±0.67	0.83±1 [0]	1.1±0.84 [0]	1.6±1.3 [0,1]	2.3±2.2 [0,1,2]
DDA	phasic (μS)	3±2.4	2.6±2.1	2.8±2.4	3.2±2.7	3.6±2.6 [0,1,2]
	tonic (μS)	5±4.4	4.6±4.3	5±4.5	5.4±4.7	5±4.2
	dphEDA (μS)	-0.068±0.14	0.06±0.14 [0]	0.086±0.13 [0]	0.18±0.19 [0,1,2]	0.31±0.33 [0,1,2]
	phasic driver (a.u.)	1.1±1.4	0.79±0.82	1±1	1.4±1.5 [0,1]	1.9±2.3 [0,1]
DCM	phasic (μS)	-0.22±0.41	-0.11±0.4 [0]	-0.03±0.36 [0]	0.084±0.48 [0,1]	0.33±0.6 [0,1,2,3]
	tonic (μS)	8.2±5.1	7.4±5.1	7.8±5.2	8.7±5.4	8.3±4.8
	dphEDA (μS)	-0.031±0.12	0.083±0.11 [0]	0.11±0.12 [0,]	0.17±0.16 [0,1,]	0.26±0.24 [0,1,2]
	phasic driver (a.u.)	1.3±0.98	1.5±1 [0]	1.7±1 [0]	2.1±1.2 [0,1,2]	2.3±1.5 [0,1,2]
cvxEDA	phasic (μS)	1.2±2.1	1.1±2.9 [0]	1.1±2	1.1±1.4 [1,2]	1.8±2.6 [0,1,2]
	tonic (μS)	6.8±5	6.2±5.4	6.7±5	7.6±5.2	6.8±4.6
	dphEDA (μS)	-0.066±0.18	0.08±0.16 [0]	0.11±0.16 [0]	0.19±0.22 [0,1]	0.31±0.38 [0,1,2]
	phasic driver (a.u.)	2.4±6.4	3.7±7.5 [0]	4.2±6.5 [0]	4.6±3.8 [0,1,2]	6.9±9.3 [0,1,2]
sparsEDA	phasic (μS)	0.3±0.77	0.24±0.72	0.3±0.72	0.53±0.92 [1]	0.59±0.9 [0,1,2]
	tonic (μS)	7.7±5	7.1±5	7.5±5	8.2±5.3	8.1±4.7
	dphEDA (μS)	-0.071±0.12	0.07±0.24 [0]	0.11±0.19 [0]	0.19±0.19 [0,1]	0.32±0.32 [0,1,2]
	phasic driver (a.u.)	0.11±0.5	0.42±1.1 [0]	0.42±0.95 [0]	1±1.6 [0,1,2]	1.9±3.1 [0,1,2]
MTVSymp		0.097±0.17	0.23±0.25 [0]	0.32±0.26 [0,1]	0.5±0.38 [0,1,2]	0.66±0.45 [0,1,2]

Superscript indices indicate which levels show significant difference with that level for that approach (Dunn's test, $p < 0.05$)

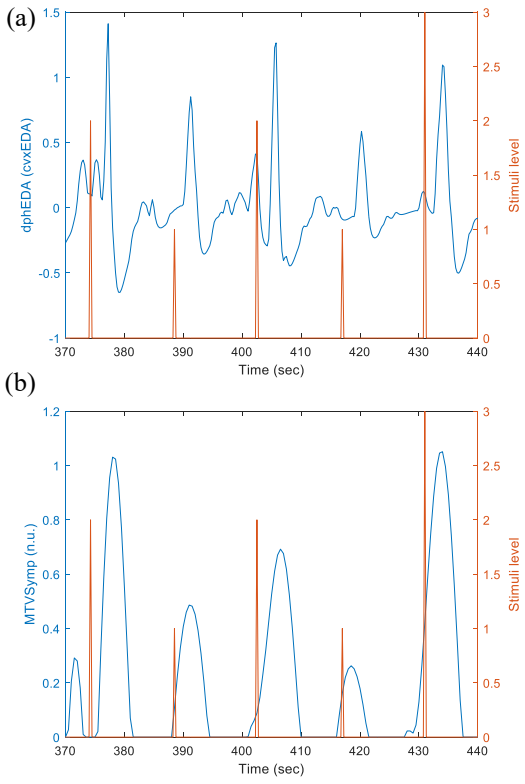


Fig. 3. Segment of (a) dphEDA of cvxEDA and (b) MTVSymp in response to electric stimulation

CDA capture the lower contour of the raw EDA signal. The tonic component obtained with DCM includes higher frequencies (more fluctuations) when compared to other approaches.

Table 1 shows the significant differences encountered between levels of stimulation (0 through 4). The phasic components obtained using DCA, DCM, and cvxEDA exhibited significant differences between level 1 and level 0. Only DCM exhibited differences between level 2 and level 0. In level 3, sparsEDA exhibited significant differences to level 1, whilst DCM exhibited differences to levels 0 and 1, and CDA and cvxEDA exhibited differences to levels 1 and 2. In level 4, CDA, DDA, cvxEDA, and sparsEDA exhibited significant differences to levels 0, 1, and 2. DCM exhibited significant differences to all the lower levels of stimulation (0 through 3). All phasic drivers' estimations showed significant differences between level 0 and all other levels. In level 3, all phasic drivers exhibited significant differences from level 0 and level 1, while DCM, cvxEDA, and sparsEDA also showed significant difference to level 2. Phasic drivers with level 4 stimulation were significantly different from level 0 through 2. No tonic component exhibited significant differences between stimulation levels.

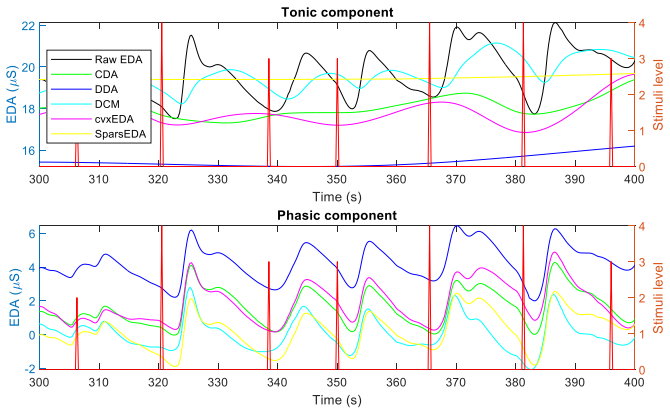


Fig. 4. Tonic (top), and phasic (bottom) components using CDA, DDA, DCM, cvxEDA, and sparsEDA

TABLE II
CLASSIFICATION PERFORMANCE USING PHASIC COMPONENTS ONLY, PHASIC COMPONENT AND TVSYMP, PHASIC COMPONENT AND MTVSYMP, PHASIC COMPONENT AND DPHEDA, AND ALL FEATURES

	Phasic Only	*+TVSymp*	*+MTVSymp*	*+dPh*	*+All BACC / M-RMSE / M-MAE*		
	B-ACC	*B-ACC*	*B-ACC*	*B-ACC*	*B-ACC*	*M-RMSE*	*M-MAE*
CDA	0.289	0.345	0.388	0.397	0.422	1.317	0.910
CDA^{PD}	0.341	0.368	0.386	0.378	0.416	1.346	0.931
DDA	0.309	0.350	0.364	0.333	0.416	1.339	0.921
DDA^{PD}	0.220	0.310	0.368	0.380	0.432	1.354	0.924
cvxEDA	0.315	0.346	0.386	0.386	0.441	1.377	0.934
cvxEDA^{PD}	0.335	0.373	0.388	0.415	0.430	1.337	0.908
sparsEDA	0.309	0.343	0.342	0.406	0.411	1.279	0.888
sparseEDA^{PD}	0.328	0.345	0.400	0.393	0.429	1.310	0.883
DCM	0.319	0.349	0.371	0.386	0.423	1.312	0.896
DCM^{PD}	0.292	0.319	0.344	0.352	0.402	1.449	0.997

B-ACC: Balanced accuracy, PD: phasic driver

Table 2 includes the measures of performance for the five-class classification, comparing results of R-SVM with different feature combinations. All approaches achieved a balanced accuracy below 35% using phasic components only, although the decomposition approach that achieved the highest balanced accuracy was CDA phasic driver (34.1%). When TVSymp and MTVSymp were included in the classification algorithm, the balanced accuracy increased to 37.3% and 40.0%, respectively. With dphEDA, the balanced accuracy was up to 41.5% with cvxEDA phasic driver. When all the features were included in the classification model, the balanced accuracy achieved its highest value of 44.1%, using cvxEDA components. Also, M-RMSE and M-MAE were the lowest among the methods with 1.279 and 0.883 using all features with SparsEDA and sparsEDA phasic driver components, respectively.

TABLE III
COMPARISON OF MLP AND R-SVM WITH ZERO-R CLASSIFIER WITH ALL FEATURES (PHASIC, DPHEDA, TVSYMP, AND MTVSYMP)

	MLP			*R-SVM*		
	B-ACC	*M-RMSE*	*M-MAE*	*B-ACC*	*M-RMSE*	*M-MAE*
Zero-R	0.190	2.103	1.764	0.190	2.103	1.764
CDA	0.421	1.297	0.892	0.422	1.317	0.910
CDA^{PD}	0.427	1.324	0.903	0.416	1.346	0.931
DDA	0.418	1.333	0.918	0.416	1.339	0.921
DDA^{PD}	0.418	1.312	0.904	0.432	1.354	0.924
cvxEDA	0.423	1.380	0.944	0.441	1.377	0.934
cvxEDA^{PD}	0.411	1.311	0.901	0.430	1.337	0.908
sparsEDA	0.420	1.280	0.885	0.411	1.279	0.888
sparsEDA^{PD}	0.427	1.289	0.877	0.429	1.310	0.883
DCM	0.434	1.328	0.907	0.423	1.312	0.896
DCM^{PD}	0.397	1.362	0.956	0.402	1.449	0.997

B-ACC: Balanced accuracy, PD: phasic driver

Table 3 shows the comparison of MLP and R-SVM with a Zero-R classifier, which shows what is the minimum expected performance for the evaluation metric. Both MLP and R-SVM exhibited the minimum balanced accuracy of 39.7%. Zero-R classifier showed 19%, 2.103, and 1.764 for balanced accuracy, M-RMSE, and M-MAE, respectively. Table 4 shows the confusion matrix for the model that showed the highest balanced accuracy, which was found to be the R-SVM algorithm with the phasic components from the cvxEDA decomposition, dphEDA, TVSymp, and MTVSymp. Level 0 was the most accurately classified (72.8% hits), followed by Level 1 and Level 4 (44% and 42% hits respectively). Levels 2 and 3 were correctly classified in 39% and 24% of the instances.

TABLE IV
CONFUSION MATRIX OF THE BEST MODEL FOR THE FIVE-CLASS CLASSIFICATION (CVXEDA)

		Predicted				
		Level 0	*Level 1*	*Level 2*	*Level 3*	*Level 4*
True	*Level 0*	0.72	0.17	0.06	0.03	0.02
	Level 1	0.18	0.44	0.22	0.07	0.10
	Level 2	0.12	0.26	0.39	0.14	0.09
	Level 3	0.11	0.15	0.22	0.24	0.27
	Level 4	0.08	0.14	0.19	0.18	0.42

TABLE V
RESULTS FOR CLASSIFICATION INTO LOW, MEDIUM, AND HIGH LEVELS OF PAIN STIMULATION WITH ALL FEATURES (PHASIC, DPHEDA, TVSYMP, AND MTVSYMP)

		B-Acc	M-RMSE	M-MAE
	Zero-R	0.331	1.214	0.967
R-SVM	CDA	0.613	0.787	0.474
	CDA^{PD}	0.619	0.766	0.455
	DDA	0.620	0.780	0.461
	DDA^{PD}	0.601	0.793	0.481
	cvxEDA	0.584	0.767	0.459
	cvxEDA^{PD}	0.597	0.781	0.480
	sparsEDA	0.631	0.756	0.442
	sparsEDA^{PD}	0.624	0.740	0.439
	DCM	0.601	0.789	0.480
	DCM^{PD}	0.614	0.786	0.471

sparsEDA		Predicted		
		low	*medium*	*high*
Known	*low*	81%	12%	7%
	medium	28%	53%	19%
	high	15%	30%	55%

B-ACC: Balanced accuracy, PD: phasic driver

Considering the misclassifications and overlap between levels of stimulation (especially between pain levels 1 and 2, 2 and 3, and 3 and 4), we tried a classification approach using

TABLE VI
HIGH-LEVEL PAIN CLASSIFICATION PERFORMANCE USING PHASIC COMPONENTS ONLY, PHASIC COMPONENT AND TVSYMP, PHASIC COMPONENT AND MTVSYMP, PHASIC COMPONENT AND DPHEDA, AND ALL FEATURES

	Phasic Only	*+TVSymp*	*+MTVSymp*	*+dPh*		*+All*	
	ACC	*ACC*	*ACC*	*ACC*	*ACC*	*Sensitivity*	*Specificity*
CDA	0.787	0.822	0.849	0.856	0.848	0.814	0.865
CDA^{PD}	0.788	0.811	0.832	0.856	0.850	0.833	0.868
DDA	0.800	0.824	0.849	0.841	0.843	0.820	0.855
DDA^{PD}	0.574	0.752	0.817	0.829	0.843	0.802	0.884
cvxEDA	0.804	0.822	0.868	0.860	0.870	0.830	0.890
cvxEDA^{PD}	0.794	0.789	0.838	0.834	0.856	0.861	0.852
sparsEDA	0.783	0.799	0.837	0.868	0.856	0.789	0.890
sparsEDA^{PD}	0.815	0.779	0.825	0.860	0.856	0.854	0.858
DCM	0.788	0.832	0.853	0.850	0.853	0.820	0.869
DCM^{PD}	0.653	0.816	0.823	0.854	0.842	0.845	0.841

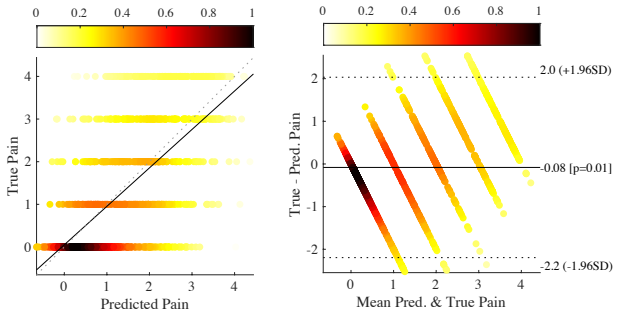
ACC: accuracy, *PD: phasic driver*

"low" (stimulation levels 0 and 1), "medium" (stimulation level 2), and "high" (stimulation levels 3 and 4) pain levels. Again, R-SVM was tried, for all phasic components, including dphEDA and MTVSymp. Table 5 shows the results of the three pain-level classification, along with the resulting confusion matrix for the best model, based on balanced accuracy. The resulting models achieved balanced accuracies around 60%. Zero-R classifier showed 33.1%, 1.214, and 0.967 for balanced accuracy, M-RMSE, and M-MAE, respectively. The highest balanced accuracy of 63% was achieved by using sparse phasic components, dphEDA, TVSymp, and MTVSymp. The lowest M-RMSE of 0.740 and M-MAE 0.439 were obtained by using the sparse phasic driver, dphEDA, and MTVSymp.

TABLE VII
COMPARISON OF MLP AND R-SVM WITH ZERO-R REGRESSORS WITH ALL FEATURES (PHASIC, DPHEDA, TVSYMP, AND MTVSYMP)

	MLP			*R-SVM*		
	R^2	*M-RMSE*	*M-MAE*	R^2	*M-RMSE*	*M-MAE*
Zero-R	-0.253	1.208	1.204	-0.253	1.208	1.204
CDA	0.459	1.116	0.880	0.441	1.170	0.916
CDA^{PD}	0.461	1.121	0.894	0.422	1.188	0.929
DDA	0.446	1.142	0.914	0.425	1.176	0.923
DDA^{PD}	0.380	1.165	0.937	0.425	1.183	0.936
cvxEDA	0.389	1.207	0.930	0.409	1.222	0.970
cvxEDA^{PD}	0.433	1.144	0.900	0.411	1.192	0.948
sparsEDA	0.464	1.117	0.889	0.441	1.155	0.941
sparsEDA^{PD}	0.477	1.129	0.885	0.427	1.178	0.925
DCM	0.459	1.123	0.875	0.457	1.150	0.897
DCM^{PD}	0.439	1.131	0.913	0.367	1.246	0.980

R^2: coefficient of determination, PD: phasic driver



(a) (b)

Fig. 5. An example of regression performance of MLP using sparsEDA phasic driver, dPhEDA, TVSymp, and MTVSymp. (a) Normalized density correlation plot with the regression line (solid line) and a Pearson correlation coefficient of 1 (dashed line). (b) Normalized density Bland-Altman plot with a mean difference of -0.08 with the limit of agreement of 95% (dashed lines).

We also trained R-SVM models to discriminate high levels of pain from no pain, as shown in Table 6. With phasic components, accuracies were between 57.4 – 81.5%. With dPhEDA or MTVSymp features, accuracies were higher than with TVSymp. All features together showed higher accuracy than using only phasic features, but it showed similar results with either TVSymp or MTVSymp only. The highest accuracy, sensitivity, and specificity were 87%, 83%, and 89%, respectively, with the cvxEDA decomposition method. Specificity was higher than sensitivity for all decomposition methods except for DCM phasic driver.

TABLE VIII
PERFORMANCE OF MULTI-LAYER PERCEPTRON REGRESSORS USING PHASIC COMPONENTS ONLY, PHASIC COMPONENT AND TVSYMP, PHASIC COMPONENT AND MTVSYMP, PHASIC COMPONENT AND DPHEDA, AND ALL FEATURES

	Phasic Only	*+TVSymp*	*+MTVSymp*	*+dPh*		*+All*	
	R^2	R^2	R^2	R^2	R^2	*M-RMSE*	*M-MAE*
CDA	0.225	0.282	0.317	0.451	0.459	1.116	0.880
CDA^{PD}	0.274	0.299	0.390	0.439	0.461	1.121	0.894
DDA	0.207	0.257	0.328	0.404	0.446	1.142	0.914
DDA^{PD}	-0.249	0.120	0.312	0.410	0.380	1.165	0.937
cvxEDA	0.208	0.277	0.348	0.396	0.389	1.207	0.930
cvxEDA^{PD}	0.258	0.283	0.385	0.347	0.433	1.144	0.900
sparsEDA	0.166	0.246	0.341	0.454	0.464	1.117	0.889
sparsEDA^{PD}	0.215	0.299	0.355	0.468	0.477	1.129	0.885
DCM	0.307	0.319	0.356	0.443	0.459	1.123	0.875
DCM^{PD}	0.006	0.234	0.335	0.383	0.439	1.131	0.913

R^2: coefficient of determination, PD: phasic driver

Table 7 shows the regressor performance of MLP and R-SVM with Zero-R regressors. MLP showed lower M-RMSE and M-MAE with higher coefficient of determination except when using cvxEDA. The phasic driver of sparsEDA showed the highest coefficient of determination, 0.477, as shown in Figure 5. The lowest M-RMSE and M-MAE were 1.116 for CDA and 0.875 for DCM from MLP regressors, respectively. As MLP regressors outperformed SVM-R regressors, we benchmarked MLP results for each EDA feature in Table 8. Coefficients of determination ranged between -0.249 - 0.307 with only the phasic feature. With dPhEDA features, the coefficient of determination was higher than with TVSymp or MTVSymp as it ranged from 0.34 - 0.47. With all features, the coefficient of determination, M-RMSE, and M-MAE ranged between 0.38 - 0.48, 1.11 - 1.21, and 0.87 - 0.93, respectively.

IV. Discussion

We have tested two features based on the differential characteristics of EDA, along with five different techniques for decomposing EDA data into tonic and phasic components. We have used the features and components to train machine learning models for pain stimulation level classification (0 through 4). The balanced accuracy of the models increased from about 30% to about 44% when the differential characteristics of EDA were included into the classification model, and the lowest macro-averaged RMSE and MAE (M-RMSE and M-MAE) were 1.279 and 0.883, respectively. Please note that the baseline classifier (Zero-R) had a balanced accuracy of 19%, an M-RMSE of 2.103, and an M-MAE of 1.764. Furthermore, we were able to classify low, medium, and high pain stimulation levels with a balanced accuracy of 63%, M-RMSE of 0.756, and M-MAE of 0.442. Also, we trained binary classifiers for detecting high levels of pain versus no pain and obtained up to 87% accuracy, 83% sensitivity, and 89% specificity. In addition to classification, we performed regression analysis and obtained up to 0.477 of coefficient of determination with 1.129 for M-RMSE and 0.885 for M-MAE. Features based on analysis of the differential characteristics of EDA have shown promising results for objective pain assessment.

As has been shown before, different approaches for obtaining tonic and phasic components of EDA provide largely different estimations [9]. We found that DCM provides a phasic component estimation slightly more sensitive to electric pain stimulation level. The traditional DCM approach has as its main limitation the computational time required (10 to 100 s per minute of data) [42]. We have used a faster implementation of DCM, although it is still slow for real-time applications (1.21±0.33 s/min). The cvxEDA and sparsEDA algorithms are faster approaches, compared to CDA, DDA, and DCM. Those approaches only exhibited significant consistent differences between the highest and the lowest levels of pain. CDA, and more consistently DCM, exhibited more resolution in their differences between levels of stimulation. However, cvxEDA outperformed other methods for the five-class classification and the high-level pain classification while sparsEDA showed higher balanced accuracy for the three-level pain classification and higher coefficient of determination for the regression than other approaches.

We found that our differential features can improve the performance for pain quantification, as compared with using only traditional EDA features (phasic components and TVSymp). However, more variables should be considered including gender, genetic background, and ethnicity, as pain is subjective perception affected by biological, sociocultural, and psychological factors [63]. Moreover, adaptation may have affected our results as the inter-stimulus interval (~30 seconds) was repeated. Further examination of these issues will be required in future studies. Other information also can be considered such as heart rate (obtained using photoplethysmography, for example) to increase classification and regression performance. Finally, the small number of subjects in our dataset and given that we have not further evaluated our models with an external dataset, limits the generalizability of our approach. To address this issue, our future studies will require further validation using data from different scenarios (e.g., different types of pain, stress, chronic pain, etc.).

V. Conclusion

Information about pain stimulation resides mostly in the phasic component of EDA. We have found that relevant information can be found from the differential characteristics of the EDA. Incorporating those features, namely the derivative of the phasic component of EDA (dphEDA) and the modified time-varying spectral components of EDA (MTVSymp) when combined with machine-learning, the best performance was achieved for classification and regression of pain stimulation levels. The phasic components of EDA along with MTVSymp are promising approaches for developing an objective measure of different pain levels.

Acknowledgment

Authors H.F.P-Q. and Y.K. equally contributed in the authorship of this paper.

References

[1] I. S. Thong, M. P. Jensen, J. Miró, and G. Tan, "The validity of pain intensity measures: what do the NRS, VAS, VRS, and FPS-R measure?," *Scandinavian journal of pain*, vol. 18, no. 1, pp. 99–107, 2018.

[2] M. P. Jensen and P. Karoly, "Self-report scales and procedures for assessing pain in adults.," 2011.

[3] M. J. Hjermstad *et al.*, "Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review," *Journal of pain and symptom management*, vol. 41, no. 6, pp. 1073–1093, 2011.

[4] M. L. Hill and K. D. Craig, "Detecting deception in pain expressions: the structure of genuine and deceptive facial displays," *Pain*, vol. 98, no. 1–2, pp. 135–144, 2002.

[5] A.-C. Larochette, C. T. Chambers, and K. D. Craig, "Genuine, suppressed and faked facial expressions of pain in children," *Pain*, vol. 126, no. 1–3, pp. 64–71, 2006.

[6] H. Nahman-Averbuch and R. C. Coghill, "Pain-autonomic relationships: implications for experimental design and the search for an 'objective marker' for pain," *PAIN*, vol. 158, no. 11, p. 2064, Nov. 2017, doi: 10.1097/j.pain.0000000000001035.

[7] M. M. Nickel, E. S. May, L. Tiemann, M. Postorino, S. Ta Dinh, and M. Ploner, "Autonomic responses to tonic pain are more closely

related to stimulus intensity than to pain intensity," *Pain*, vol. 158, no. 11, pp. 2129–2136, Nov. 2017, doi: 10.1097/j.pain.0000000000001010.

[8] M. M. Nickel *et al.*, "Brain oscillations differentially encode noxious stimulus intensity and pain intensity," *Neuroimage*, vol. 148, pp. 141–147, 01 2017, doi: 10.1016/j.neuroimage.2017.01.011.

[9] H. F. Posada-Quintero and K. H. Chon, "Innovations in Electrodermal Activity Data Collection and Signal Processing: A Systematic Review," *Sensors*, vol. 20, no. 2, p. 479, Jan. 2020, doi: 10.3390/s20020479.

[10] D. R. Bach, "Sympathetic nerve activity can be estimated from skin conductance responses - a comment on Henderson et al. (2012)," *Neuroimage*, vol. 84, pp. 122–123, Jan. 2014, doi: 10.1016/j.neuroimage.2013.08.030.

[11] M. E. Dawson, A. M. Schell, and D. L. Filion, "Chapter 7: the electrodermal system," *Handbook of psychophysiology*, pp. 159–181, 2007.

[12] W. Boucsein *et al.*, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, Aug. 2012, doi: 10.1111/j.1469-8986.2012.01384.x.

[13] G. Bohlin, "Delayed habituation of the electrodermal orienting response as a function of increased level of arousal," *Psychophysiology*, vol. 13, no. 4, pp. 345–351, Jul. 1976, doi: 10.1111/j.1469-8986.1976.tb03088.x.

[14] S. Bruehl and O. Y. Chung, "Interactions between the cardiovascular and pain regulatory systems: an updated review of mechanisms and possible alterations in chronic pain," *Neuroscience & Biobehavioral Reviews*, vol. 28, no. 4, pp. 395–414, 2004.

[15] J. Koenig, M. N. Jarczok, R. J. Ellis, T. K. Hillecke, and J. F. Thayer, "Heart rate variability and experimentally induced pain in healthy adults: a systematic review," *European Journal of Pain*, vol. 18, no. 3, pp. 301–314, 2014.

[16] S. Duschek, I. Mück, and G. A. Reyes Del Paso, "Relationship between baroreceptor cardiac reflex sensitivity and pain experience in normotensive individuals," *Int J Psychophysiol*, vol. 65, no. 3, pp. 193–200, Sep. 2007, doi: 10.1016/j.ijpsycho.2007.03.012.

[17] M. L. Loggia, M. Juneau, and M. C. Bushnell, "Autonomic responses to heat pain: Heart rate, skin conductance, and their relation to verbal ratings and stimulus intensity," *Pain*, vol. 152, no. 3, pp. 592–598, Mar. 2011, doi: 10.1016/j.pain.2010.11.032.

[18] H. Nahman-Averbuch *et al.*, "Associations between autonomic dysfunction and pain in chemotherapy-induced polyneuropathy," *Eur J Pain*, vol. 18, no. 1, pp. 47–55, Jan. 2014, doi: 10.1002/j.1532-2149.2013.00349.x.

[19] R. Treister, M. Kliger, G. Zuckerman, I. Goor Aryeh, and E. Eisenberg, "Differentiating between heat pain intensities: the combined effect of multiple autonomic parameters," *Pain*, vol. 153, no. 9, pp. 1807–1814, Sep. 2012, doi: 10.1016/j.pain.2012.04.008.

[20] J. T. Cacioppo, L. G. Tassinary, G. Berntson, and & 0 more, *Handbook of Psychophysiology*, 3 edition. Cambridge England ; New York: Cambridge University Press, 2007.

[21] "Techniques in Psychophysiology. By I. Martin and P. H. Venables. (Pp. 669; illustrated; £19.50.) John Wiley & Sons: Chichester. 1980.," *Psychological Medicine*, vol. 10, no. 04, pp. 806–806, 1980, doi: 10.1017/S0033291700055343.

[22] N. S. Greenfield and R. A. Sternbach, *Handbook of psychophysiology*, vol. xii. Oxford, England: Holt, Rinehart & Winston, 1972.

[23] K. Kobielak, E. Kandyba, and Y. Leung, "Skin and Skin Appendage Regeneration," in *Translational Regenerative Medicine*, Elsevier, 2015, pp. 269–292.

[24] P. H. Ellaway, A. Kuppuswamy, A. Nicotra, and C. J. Mathias, "Sweat production and the sympathetic skin response: Improving the clinical assessment of autonomic function," *Autonomic Neuroscience*, vol. 155, no. 1–2, pp. 109–114, Jun. 2010, doi: 10.1016/j.autneu.2010.01.008.

[25] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *J Neurosci Methods*, vol. 190, no. 1–5, pp. 80–91, Jun. 2010, doi: 10.1016/j.jneumeth.2010.04.028.

[26] D. J. Clark, S. A. Chatterjee, T. E. McGuirk, E. C. Porges, E. J. Fox, and C. K. Balasubramanian, "Sympathetic nervous system activity measured by skin conductance quantifies the challenge of walking adaptability tasks after stroke," *Gait & posture*, vol. 60, pp. 148–153, 2018.

[27] H. F. Posada-Quintero *et al.*, "Using electrodermal activity to validate multilevel pain stimulation in healthy volunteers evoked by thermal grills," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 319, no. 3, pp. R366–R375, 2020.

[28] S. Sugimine, S. Saito, and T. Takazawa, "Normalized skin conductance level could differentiate physical pain stimuli from other sympathetic stimuli," *Scientific reports*, vol. 10, no. 1, pp. 1–12, 2020.

[29] D. Lopez-Martinez and R. Picard, "Continuous pain intensity estimation from autonomic signals with recurrent neural networks," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 5624–5627.

[30] W. Boucsein, *Electrodermal Activity*. Springer Science & Business Media, 2012.

[31] D. R. Bach, J. Daunizeau, N. Kuelzow, K. J. Friston, and R. J. Dolan, "Dynamic causal modeling of spontaneous fluctuations in skin conductance," *Psychophysiology*, vol. 48, no. 2, pp. 252–257, Feb. 2011, doi: 10.1111/j.1469-8986.2010.01052.x.

[32] A. Greco, G. Valenza, A. Lanata, E. Scilingo, and L. Citi, "cvxEDA: a Convex Optimization Approach to Electrodermal Activity Processing," *IEEE Trans Biomed Eng*, Aug. 2015, doi: 10.1109/TBME.2015.2474131.

[33] F. Hernando-Gallego, D. Luengo, and A. Artes-Rodriguez, "Feature Extraction of Galvanic Skin Responses by Non-Negative Sparse Deconvolution," *IEEE J Biomed Health Inform*, Dec. 2017, doi: 10.1109/JBHI.2017.2780252.

[34] M. R. Amin and R. T. Faghih, "Identification of Sympathetic Nervous System Activation from Skin Conductance: A Sparse Decomposition Approach with Physiological Priors," *IEEE Transactions on Biomedical Engineering*, 2020.

[35] D. S. Wickramasuriya and R. T. Faghih, "A Marked Point Process Filtering Approach for Tracking Sympathetic Arousal From Skin Conductance," *IEEE Access*, vol. 8, pp. 68499–68513, 2020.

[36] D. S. Wickramasuriya and R. T. Faghih, "A mixed filter algorithm for sympathetic arousal tracking from skin conductance and heart rate measurements in Pavlovian fear conditioning," *Plos one*, vol. 15, no. 4, p. e0231659, 2020.

[37] H. F. Posada-Quintero, J. P. Florian, A. D. Orjuela-Cañón, T. Aljama-Corrales, S. Charleston-Villalobos, and K. H. Chon, "Power Spectral Density Analysis of Electrodermal Activity for Sympathetic Function Assessment," *Ann Biomed Eng*, vol. 44, no. 10, pp. 3124–3135, Apr. 2016, doi: 10.1007/s10439-016-1606-6.

[38] H. F. Posada-Quintero, J. P. Florian, Á. D. Orjuela-Cañón, and K. H. Chon, "Highly sensitive index of sympathetic activity based on time-frequency spectral analysis of electrodermal activity," *American Journal of Physiology - Regulatory, Integrative and Comparative Physiology*, vol. 311, no. 3, pp. R582–R591, Sep. 2016, doi: 10.1152/ajpregu.00180.2016.

[39] H. F. Posada-Quintero, T. Dimitrov, A. Moutran, S. Park, and K. H. Chon, "Analysis of Reproducibility of Noninvasive Measures of Sympathetic Autonomic Control Based on Electrodermal Activity and Heart Rate Variability," *IEEE Access*, vol. 7, pp. 22523–22531, 2019, doi: 10.1109/ACCESS.2019.2899485.

[40] Y. Kong, H. F. Posada-Quintero, and K. H. Chon, "Pain Detection using a Smartphone in Real Time," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, pp. 4526–4529.

[41] "257 - Safe Use of Electrical Stimulators | BIOPAC," *BIOPAC Systems, Inc.* https://www.biopac.com/application-note/stimulation-safe-use-of-electrical-stimulators/ (accessed Dec. 26, 2019).

[42] D. R. Bach and M. Staib, "A matching pursuit algorithm for inferring tonic sympathetic arousal from spontaneous skin conductance fluctuations," *Psychophysiology*, vol. 52, no. 8, pp. 1106–1112, Aug. 2015, doi: 10.1111/psyp.12434.

[43] "Central Differences." http://www.holoborodko.com/pavel/numerical-methods/numerical-derivative/central-differences/ (accessed Jun. 23, 2020).

[44] K. H. Chon, S. Dash, and K. Ju, "Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation," *IEEE Trans Biomed Eng*, vol. 56, no. 8, pp. 2054–2063, Aug. 2009, doi: 10.1109/TBME.2009.2019766.

[45] H. Wang, K. Siu, K. Ju, and K. H. Chon, "A high resolution approach to estimating time-frequency spectra and their amplitudes," *Ann Biomed Eng*, vol. 34, no. 2, pp. 326–338, Feb. 2006, doi: 10.1007/s10439-005-9035-y.

[46] H. Gasquet and A. J. Wootton, "Variable-frequency complex demodulation technique for extracting amplitude and phase

information," *Review of Scientific Instruments*, vol. 68, no. 1, pp. 1111–1114, Jan. 1997, doi: 10.1063/1.1147748.

[47] A. Monti, C. Médigue, and L. Mangin, "Instantaneous parameter estimation in cardiovascular time series by harmonic and time-frequency analysis," *IEEE Trans Biomed Eng*, vol. 49, no. 12 Pt 2, pp. 1547–1556, Dec. 2002, doi: 10.1109/TBME.2002.805478.

[48] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, Mar. 1998, doi: 10.1098/rspa.1998.0193.

[49] F. J. Massey Jr, "The Kolmogorov-Smirnov test for goodness of fit," *Journal of the American statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.

[50] L. H. Miller, "Table of percentage points of Kolmogorov statistics," *Journal of the American Statistical Association*, vol. 51, no. 273, pp. 111–121, 1956.

[51] J. Wang, W. W. Tsang, and G. Marsaglia, "Evaluating Kolmogorov's distribution," *Journal of Statistical Software*, vol. 8, no. 18, 2003.

[52] G. Cardillo, *Dunn Test: a procedure for multiple, not parametric, comparisons*. MATLAB Central, MathWorks, Natick, MA, 2006.

[53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[54] E. Alpaydin, *Introduction to Machine Learning*. MIT Press, 2020.

[55] R. Fletcher, *Practical methods of optimization*. John Wiley & Sons, 2013.

[56] S. Saeb, L. Lonini, A. Jayaraman, D. C. Mohr, and K. P. Kording, "The need to approximate the use-case in clinical machine learning," *Gigascience*, vol. 6, no. 5, May 2017, doi: 10.1093/gigascience/gix019.

[57] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[58] H. M. Nguyen, E. W. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 3, no. 1, pp. 4–21, 2011.

[59] G. Lemaître, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 559–563, 2017.

[60] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[61] S. Baccianella, A. Esuli, and F. Sebastiani, "Evaluation measures for ordinal regression," in *2009 Ninth international conference on intelligent systems design and applications*, 2009, pp. 283–287.

[62] L. Gaudette and N. Japkowicz, "Evaluation methods for ordinal classification," in *Canadian conference on artificial intelligence*, 2009, pp. 207–210.

[63] R. B. Fillingim, "Individual differences in pain responses," *Current rheumatology reports*, vol. 7, no. 5, pp. 342–347, 2005.