Received: 30 August 2021 | Revised: 15 April 2022 | Accepted: 2 May 2022

DOI: 10.4218/etrij.2021-0299

ETRI Journal WILEY

ORIGINAL ARTICLE

Automated detection of panic disorder based on multimodal physiological signals using machine learning

Eun Hye Jang[1] | **Kwan Woo Choi**[2] | **Ah Young Kim**[1] | **Han Young Yu**[3] | **Hong Jin Jeon**[4] | **Sangwon Byun**[5]

[1]Welfare and Medical ICT Research Department, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea

[2]Department of Psychiatry, Korea University, Seoul, Republic of Korea

[3]Industry and IoT Intelligence Research Department, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea

[4]Department of Psychiatry in the Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[5]Department of Electronics Engineering, Incheon National University, Incheon, Republic of Korea

Correspondence
Sangwon Byun, Department of Electronics Engineering, Incheon National University, Incheon, Republic of Korea.
Email: swbyun@inu.ac.kr

Funding information
This work was supported by the Electronics and Telecommunications Research Institute (ETRI)'s internal funds, Development of Digital Biopsy Core Technology for high-precision Diagnosis and Therapy of Senile Disease (21YR2500), Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No. 2015-0-00062), and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1F1A1049236).

Abstract

We tested the feasibility of automated discrimination of patients with panic disorder (PD) from healthy controls (HCs) based on multimodal physiological responses using machine learning. Electrocardiogram (ECG), electrodermal activity (EDA), respiration (RESP), and peripheral temperature (PT) of the participants were measured during three experimental phases: rest, stress, and recovery. Eleven physiological features were extracted from each phase and used as input data. Logistic regression (LoR), *k*-nearest neighbor (KNN), support vector machine (SVM), random forest (RF), and multilayer perceptron (MLP) algorithms were implemented with nested cross-validation. Linear regression analysis showed that ECG and PT features obtained in the stress and recovery phases were significant predictors of PD. We achieved the highest accuracy (75.61%) with MLP using all 33 features. With the exception of MLP, applying the significant predictors led to a higher accuracy than using 24 ECG features. These results suggest that combining multimodal physiological signals measured during various states of autonomic arousal has the potential to differentiate patients with PD from HCs.

KEYWORDS
anxiety disorder, autonomic nervous system (ANS) response, deep learning, electrocardiogram (ECG), heart rate variability (HRV), machine learning, mental stress task, multimodal, panic disorder, physiological signals

1 | INTRODUCTION

Mental disorders are related to the dysfunctions of the autonomic nervous system (ANS) [1–4]. Physiological signals, which have been used to assess ANS activity, have attracted significant interest in the study of mental disorders. For example, recent studies have used physiological signals for the automated diagnosis of mental disorders. The detection of major depressive disorders based on machine learning has been performed using physiological features extracted from electrocardiogram (ECG) and electrodermal activity (EDA) signals [1–3]. Valenza and others [4] used heart rate variability (HRV) parameters to evaluate depressive states and predict mood changes in patients with bipolar disorder. These studies focused on the diagnosis of depressive disorders.

Panic disorder (PD) is one of the most common anxiety disorders and is characterized by recurrent and unexpected panic attacks [5]. It can negatively affect the personal and social lives of patients [6]. Although PD is diagnosed following the guidelines provided by the Diagnostic and Statistical Manual of Mental Disorders, the diagnosis primarily depends on clinical interviews and subjective reports of symptoms by the patients [7–9]. In addition, controlled studies on the psychological aspects of PD have not been conducted extensively owing to methodological limitations [10]. Hence, researchers have been working to develop more reliable methods for diagnosing PD based on quantitative data, such as physiological signals.

The clinical symptoms of PD include ANS disturbances, such as palpitations, tachycardia, sweating, shaking, dyspnea, and chest pain [11]. In particular, PD affects cardiac activity and increases the risk of cardiovascular morbidity and mortality. Patients with PD have a high baseline heart rate (HR), periods of tachycardia that coincide with panic symptoms [12, 13], and reduced resting-state HRV, reflecting a decrease in vagal output [14]. Previous studies on PD primarily focused on changes in HRV. A meta-analysis [14] of 24 studies comparing patients with PD and controls reported lower HRV and high-frequency (HF) power and higher low-frequency (LF) power and LF/HF ratio in patients with PD compared with controls [15–19]. However, some studies have demonstrated mixed results, reporting lower HF power and LF/HF ratio in patients with PD [19–22]. Kotianova and others reported lower very low frequency (VLF) power during baseline measurement and higher LF/HF ratio during response to mental tasks in patients with PD compared with controls [23].

These results suggest that it is important to include physiological features other than HRV in the investigation of abnormal changes in autonomic activity in patients with PD. EDA parameters, such as skin conductance level (SCL), reflect sympathetic nervous system activity and are sensitive to changes in clinical autonomic status [24]. Moreover, finger temperature (FT) represents peripheral physiological response to anxiety [25]. Patients with PD were found to have significantly higher HR and SCL and significantly lower FT than healthy controls (HCs) [25]. Respiration (RESP) has also been studied as a physiological indicator of stress and anxiety [26, 27], and respiration rate (RR) has been shown to increase as the intensity of stress or anxiety increases [28]. Panic is characterized by stress-induced autonomic sensations, and the effects of experimentally induced psychosocial stress on several ANS responses have been investigated. Previous studies involved the measurement of physiological signals while patients were conducting experimental tasks and demonstrated that patients with PD and generalized anxiety disorder (GAD) have higher SCL than patients with major depressive disorder [29, 30]. Therefore, subjecting patients to stress in a laboratory setting may aid the development of approaches for improving the discriminative power of methods using physiological features for detecting PD.

Based on these results of abnormal ANS activity in patients with PD, previous studies used machine learning algorithms to differentiate PD from other pathological conditions or predict treatment outcomes. For example, Na and others applied five algorithms to HRV data to discriminate PD from other anxiety disorders [31]. Lueken and others distinguished depressive comorbidity from PD using functional magnetic resonance imaging (fMRI) data and a tree ensemble classifier [32]. Sundermann and others used a support vector machine (SVM) with fMRI data to predict the response to cognitive behavioral therapy in patients with PD and agoraphobia [33]. However, the automated differentiation of patients with PD from HCs using physiological signals has not been extensively tested.

The aim of this study was to test the feasibility of the automated discrimination of patients with PD from HCs using machine learning approaches based on multimodal physiological responses. To obtain physiological responses, we used a psychophysiological profile (PPP), a technique used to evaluate autonomic arousal and quantify the level of individual stress reactivity, with the expectation that inducing multiple alterations in the ANS response would improve the performance of classifiers [34]. The PPP comprised three continuous phases: rest, stress presentation (cognitive or perceptual tasks), and recovery from stress [34–37]. During the PPP phases, four physiological signals were simultaneously measured: ECG, EDA, RESP, and peripheral temperature (PT). A

total of 33 physiological features—11 features from three phases—were extracted and used as input data. First, we statistically analyzed the extracted features to identify significant predictors of PD. Subsequently, various machine learning algorithms were applied to the input data to predict the PD and HC groups.

2 | EXPERIMENTAL METHODS

2.1 | Subjects

This study included 71 subjects [41 females, mean age ± standard deviation (SD), 42.3 ± 14.39 years], consisting of 39 HCs and 32 patients with PD. The patients were enrolled between December 2015 and January 2017 and were recruited from the outpatient clinic of the Depression Center of the Samsung Medical Center. They were diagnosed with PD by a senior psychiatrist if they met the DSM-IV criteria for PD and scored >7 points on the Panic Disorder Severity Scale (PDSS) [38]. The exclusion criteria were pregnancy, history of drug or alcohol dependence, history of head trauma, serious risk of suicide, personality disorder, severe somatic diseases, and use of long-acting medications, including fluoxetine and depot neuroleptics. Using general study advertisements, healthy subjects with no history of psychiatric disease and no family history of mood disorders were recruited as the HC group. The study protocol was approved by the Ethics Committee of the Samsung Medical Center, Seoul, Korea (No. 2015-07-151), and the study was performed in accordance with the relevant guidelines. Written informed consent was obtained from all participants after providing them an explanation of the experimental procedures.

2.2 | Demographic data and clinical measures

The study duration was 12 weeks for each subject (Figure 1). Each subject completed five visits: baseline and 2, 4, 8, and 12 weeks after the initial screening. All subjects provided demographic data and completed clinical measures. The demographic data included age, sex, and years of education. The clinical measures included the Hamilton Rating Scale for Anxiety (HAM-A) [39] and PDSS [38]. These scales were evaluated at the baseline visit and at the 12-week visit. We also evaluated the subjects' body mass index (BMI), smoking habits, and alcohol consumption, which are known to be associated with changes in ANS parameters [40].

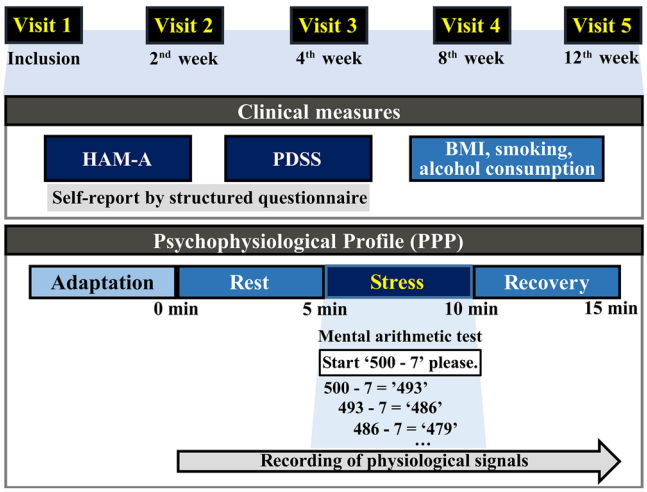

FIGURE 1 Experimental procedure

2.3 | Experimental design

The experimental protocol included PPP, one of the techniques used for assessing the influence of autonomic response on behavior. The PPP experiment is generally divided into three continuous phases: rest, stress presentation using cognitive or perceptual tasks, and recovery [29]. Each phase lasts for 5 min. For stress presentation, we used the mental arithmetic task, which involves the continuous serial subtraction of a single-digit number, 7, from a three-digit number, 500, to gradually increase the subjects' mental load. During the experiment, physiological signals, such as ECG, EDA, RESP, and PT, were simultaneously measured. The psychophysiological assessment was completed on five separate occasions during the 12-week study period. Two investigator specialists were trained to conduct the experiments. For each experiment, only one subject was examined at a time by a specialist in the clinical laboratory. The experimental procedures used in this study are shown in Figure 1.

2.4 | Physiological measures

Because ANS responses are easily affected by the subject's physiological state, which is in turn affected by factors such as time of the day, mood, and rest state, physiological signals were measured during working hours. The experiment was conducted under standard conditions in a sound-attenuated room at a temperature of 23°C and humidity of 45%–55%. Prior to the experiment, the subjects were instructed to sit comfortably in an armchair with a headrest and not to speak or move unless necessary while the devices for recording physiological signals were set and calibrated. Physiological

signals—ECG, EDA, RESP, and PT—were acquired using the ProComp Infiniti system (SA7500, Thought Technology, Canada). The sampling rate of all signals was 256 Hz. ECGs were recorded using an ECG-Flex/Pro sensor (T9306M, Thought Technology). Three ECG electrodes were placed on both forearms. The negative lead was placed on the right forearm, whereas both the positive and ground leads were placed on the left forearm. The ECG signal was passed through a 0.26-Hz to 30-Hz band-pass filter. The EDA was recorded using SC-Flex/Pro sensors (SA9309M, Thought Technology). A constant electrical voltage of 0.5 V was applied between two dry Ag/AgCl electrodes, which were strapped to the distal phalanges of the index and middle fingers of the subject's nondominant hand. The RESP signal was measured using a RESP-Flex/Pro sensor (SA9311M, Thought Technology), which was stretched around the subject's chest and measured the relative amount of chest expansion. To measure the PT, a Temp-Flex/Pro sensor (SA9310M, Thought Technology) was placed on the palmar side of the ring finger of the subject's nondominant hand. The sensor can measure a temperature range of 10°C–45°C.

2.5 | Feature extraction

Physiological signals were analyzed using the BioGraph Infiniti Software (Thought Technology), which also amplified and digitized all signals. We used the middle 3-min segment after removing the first and last 1 min from the 5-min recording of each phase (rest, stress, and recovery). The features extracted from each signal are presented in Table 1. Eight features were extracted from the ECG signals. The R-peak to R-peak intervals (RRIs), HR, SD of the RRIs (SDNN), root mean square of successive differences in the RRIs (RMSSD), and proportion of successive RRIs differing by >50 ms (pNN50) were calculated based on time-domain analysis. The SDNN reflects both sympathetic and parasympathetic activities, whereas the RMSSD and pNN50 are sensitive to parasympathetic modulation. We also evaluated power spectrum density for the RRI data using a fast Fourier transform to extract frequency-domain features (resampling method of the RRI data is not available). The absolute powers in three distinct bands, VLF (<0.04 Hz), LF (0.04 Hz–0.15 Hz), and HF (0.15 Hz–0.4 Hz), were calculated. The relative powers of the LF and HF bands were calculated as the LF/HF ratio. The SCL was extracted from the EDA signals by averaging the data points in the 3-min segment. The RR, which signified the number of breaths per minute, was estimated by counting chest movements in the RESP signals. The FT was evaluated by averaging the 3-min PT signals. Table 2 shows a sample of data from one of the patients with PD.

2.6 | Statistical analysis

A maximum of 355 datasets were potentially available (5 visits × 71 participants). However, due to participant dropout, only 344 datasets were selected for analysis. The

TABLE 2 Physiological features of a patient with PD (male, 23-year-old, first visit)

	Phases		
Features	**Rest**	**Stress**	**Recovery**
SDNN (ms)	81.74	92.74	65.05
pNN50 (%)	0.366	0.241	0.319
RMSSD (ms)	102.2	82.58	85.63
VLF (ms^2)	505.9	326.8	134.0
LF (ms^2)	704.2	1132	492.7
HF (ms^2)	2316	972.9	680.4
LF/HF ratio	0.487	1.378	0.695
HR (bpm)	52.21	60.85	51.43
SCL (μs)	0.483	5.624	3.831
RR (per minute)	14.88	12.42	12.89
FT (°C)	33.78	33.91	33.88

Abbreviation: bpm, beats per minute.

TABLE 1 Physiological features extracted from each signal

	Features		
Signals	**Rest (res)**	**Stress (str)**	**Recovery (rec)**
ECG	SDNN, pNN50, RMSSD, VLF, LF, HF, LF/HF ratio, HR	SDNN, pNN50, RMSSD, VLF, LF, HF, LF/HF ratio, HR	SDNN, pNN50, RMSSD, VLF, LF, HF, LF/HF ratio, HR
EDA	SCL	SCL	SCL
RESP	RR	RR	RR
PT	FT	FT	FT

demographic data and clinical measures of the PD and HC groups were compared using independent *t*-test and chi-square test. Linear regression analysis was used to determine the significant predictors of PD. A *p*-value of <0.05 indicated statistical significance. All analyses were performed using SPSS 22.0 (IBM, Armonk, NY, USA).

2.7 | Machine learning methods

Figure 2 shows an overview of the data processing pipeline. To classify the PD and HC groups based on physiological features, five machine learning algorithms were implemented: logistic regression (LoR), *k*-nearest neighbor (KNN), SVM with radial basis function kernel, random forest (RF), and multilayer perceptron (MLP) [1, 41, 42]. Participants with missing visits were excluded from the input data, resulting in a total of 335 datasets from 67 participants (39 controls and 28 patients). The input data were normalized: The mean was subtracted from each data point, and the result was divided by SD.

To evaluate the performance of the classifiers, we adopted a nested cross-validation (CV) approach, which can reduce optimistic bias compared with simple CV [43]. Nested CV has a subject-wise stratified 10-fold inner loop nested in a leave-one-subject-out (LOSO) outer loop. To prevent data leakage during CV, we adopted a subject-wise, rather than sample-wise, data split [44]. LOSO was chosen as the CV method because it has been demonstrated to perform well compared with other methods, such as 10-fold and 5-fold CV, when the data size is small (<100) [45]. In the outer loop, five datasets from one subject were designated as the test set. Subject-wise 10-fold CV was used for the remaining data (training and validation datasets) to optimize the hyperparameters using a grid search approach. After the validation and training datasets were split in the inner loop, we randomly under-sampled the healthy subjects in the training dataset. Most machine learning algorithms work best when each class has the same number of samples; thus, under-sampling was conducted using a subject-wise approach to match the number of subjects in the HC and PD groups in the training dataset to maximize the performance of the classifiers. The optimal parameters were those that resulted in the highest average accuracy over the 10 folds. The final optimized model was built by training with the training/validation dataset and the optimal parameters. This model was subsequently applied to the test dataset for performance evaluation. These processes were repeated 67 times in the outer loop. To further improve the estimation, the nested CV was repeated 10 times, and the results from all 10 repetitions were averaged to obtain the final evaluation of the predictive performance of each classifier. Accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) were calculated and used as performance indices.

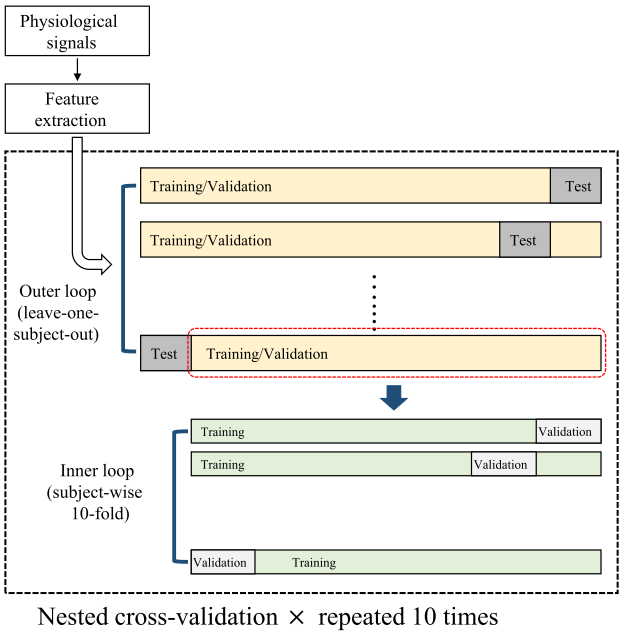

FIGURE 2 Overview of the data processing procedure. Nested cross-validation was repeated 10 times.

The following hyperparameters were optimized in the 10-fold inner loop: number of neighbors (odd numbers from 3 to 21) for KNN; C (10^{-3}, 10^{-2}, 10^{-1}, and 10^{0}) and gamma (10^{-3}, 10^{-2}, 10^{-1}, and 10^{0}) for SVM; number of trees (1000 and 2000) and maximum number of features (sqrt and none) for RF; and number of hidden layers (2, 3, and 4) for MLP. In the MLP model, the binary cross-entropy was optimized using the Adam optimization algorithm, and a 20% dropout layer was added to each of the fully connected layers before the output. Each layer contained 1024 neurons.

A total of 33 features (11 features × 3 phases) were used as the input data. Four different sets of input data were used, and their results were compared: all 33 multimodal features; six significant predictors from the linear regression analysis; 27 features, after excluding the six significant features; and 24 HRV features, after excluding the EDA, RESP, and temperature features. We computed the mutual information (MI) between the features. MI measures the dependency between two random variables [46] and can be used for feature ranking. It represents the change in the entropy of one variable from observing the other variable and is zero if and only if the two variables are independent. MI estimation was based on the KNN distances method [47, 48]. All analyses were performed using Python 3.7 and TensorFlow 2.3.

3 | EXPERIMENTAL RESULTS

3.1 | Demographic data and clinical measures

Table 3 shows the descriptive statistics of the demographic data and clinical measures of the participants in the PD and HC groups. There were no significant differences between the two groups in terms of age, sex, BMI, smoking habits, and alcohol consumption. However, the patients in the PD group had significantly more years of education than those in the HC group. We also found significant differences in the two psychiatric measures between the two groups. The HAM-A and PDSS scores were significantly higher in the PD group than in the HC group. The mean HAM-A scores in the PD and HC groups were 15.28 ± 8.16 and 2.08 ± 2.17, respectively. The mean PDSS scores in the PD and HC groups were 11.88 ± 6.64 and 0.03 ± 0.16, respectively.

3.2 | Prediction of patients with PD using physiological responses

To examine whether a combination of multimodal physiological features could predict PD, linear regression analysis with stepwise forward entry was performed (Table 4). A significant regression equation was found with $R = 0.451$ ($p < 0.001$). Six features—LF during the rest phase (res_LF), FT during the rest phase (res_FT), FT during the stress phase (str_FT), HR during the recovery phase (rec_HR), HR during the stress phase (str_HR), and pNN50 during the recovery phase (rec_pNN50)—were significant predictors and explained 20.3% of the variance in PD. PD was influenced by these physiological features in the following order: rec_HR ($\beta = 0.559$), str_HR ($\beta = -0.465$), res_FT ($\beta = -0.463$), str_FT ($\beta = 0.345$), rec_pNN50 ($\beta = -0.344$), and res_LF ($\beta = 0.269$). PD was negatively related to res_FT, str_HR, and rec_pNN50 and positively related to res_LF, str_FT,

TABLE 3 Demographic and clinical data of PD and HC groups

Factors	PD ($N = 32$)	HC ($N = 39$)	t or χ^2	p-value
Age (years ± SD)	43.47 ± 13.05	41.15 ± 15.74	0.677	0.500
Sex (male/female)	14/18	16/23	0.053	0.817
Education (years ± SD)	15.34 ± 2.31	14.05 ± 2.67	2.189	0.032
BMI (kg/m^2 ± SD)	23.72 ± 3.50	22.99 ± 2.86	0.925	0.359
Smoker/nonsmoker	6/26	6/33	0.142	0.707
Alcohol consumption (g/week ± SD)	4.77 ± 11.26	8.91 ± 9.42	1.660	0.102
HAM-A (points ± SD)	15.28 ± 8.16	2.08 ± 2.17	9.713	<0.001
PDSS (points ± SD)	11.88 ± 6.64	0.03 ± 0.16	11.163	<0.001

TABLE 4 Linear regression analysis predicting panic disorder with physiological features ($N = 337$, df = 336)

Predicted variable	Predictors	Unstandardized coefficients		Standardized coefficients			Collinearity statistics	
		B	SE	Beta	t	95% CI of B	Tolerance	VIF
Panic disorder	res_LF	0.009	0.000	0.269	4.227***	0.000 to 0.000	0.584	1.711
	res_FT	−0.156	0.041	−0.463	−3.793***	−0.237 to −0.075	0.159	6.304
	str_FT	0.127	0.045	0.345	2.799**	0.038 to 0.215	0.156	6.422
	str_HR	−0.019	0.005	−0.465	−4.138***	−0.028 to −0.010	0.187	5.342
	rec_HR	0.024	0.005	0.559	4.992***	0.015 to 0.034	0.188	5.314
	rec_pNN50	−0.027	0.005	−0.344	−5.762***	−0.037 to −0.018	0.662	1.510

Note: $R = 0.451$, $R^2 = 0.203$, adjusted $R^2 = 0.189$, $F(1,336) = 12.039$, $p < 0.001$. Durbin–Watson $d = 1.419$.
Abbreviations: CI, confidence interval; SE, standard error; VIF, variance inflation factor.
**$p < 0.01$.
***$p < 0.001$.

and rec_HR. The Durbin–Watson test showed that there was no first-order autocorrelation ($d = 1.419$). The collinearity diagnostics indicated that the variables were not affected by multicollinearity: Tolerance ranged from 0.156 to 0.662, and the variance inflation factor ranged from 1.510 to 6.422.

3.3 | Classification using machine learning algorithms

We used five machine learning algorithms—LoR, KNN, SVM, RF, and MLP—to classify the subjects into the PD or HC group using the physiological features as input data (Table 5, Figure 3). Four sets of input data were compared. First, all 33 features were used as input data. Second, six features that significantly predicted PD in the linear regression (res_LF, res_FT, str_FT, rec_HR, str_HR, and rec_pNN50) were used; these features were referred to as the significant features. The third feature

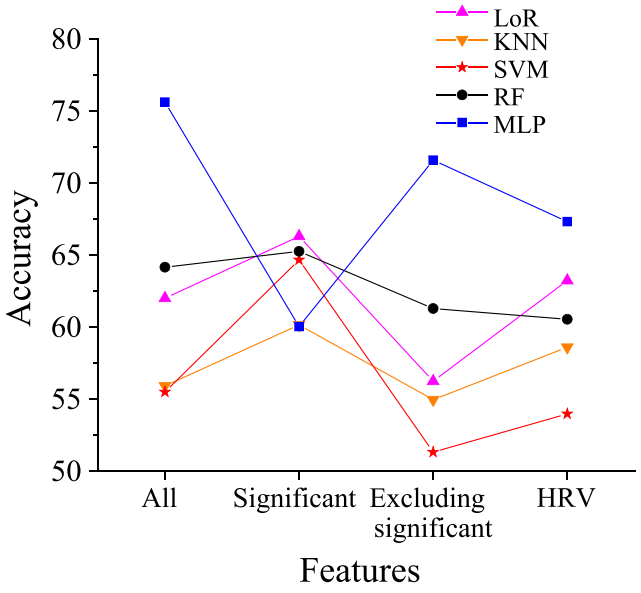

FIGURE 3 Classification accuracy for each combination of input dataset and machine learning algorithm

TABLE 5 Performance measures for the classification of PD and HC groups based on five machine learning algorithms

Algorithm/features		ACC	SEN	SPE	PPV	NPV	AUC
LoR	All	62.00	64.43	60.26	53.78	70.25	0.6670
	Significant	66.30	71.50	62.56	57.83	75.36	0.7420
	Excl. significant	56.24	60.29	53.33	48.12	65.17	0.5742
	HRV	63.22	65.86	61.33	55.02	71.45	0.6694
KNN	All	55.91	63.21	50.67	47.93	65.73	0.5990
	Significant	60.15	67.86	54.62	51.78	70.30	0.6530
	Excl. significant	54.96	55.50	54.56	46.72	63.08	0.5746
	HRV	58.60	64.93	54.05	50.34	68.28	0.6305
SVM	All	55.49	63.71	49.59	47.52	65.65	0.5961
	Significant	64.66	66.64	63.23	56.58	72.53	0.7000
	Excl. significant	51.31	60.64	44.62	43.84	61.53	0.5941
	HRV	53.97	65.64	45.59	46.30	65.22	0.6169
RF	All	64.15	61.79	65.85	56.51	70.59	0.6604
	Significant	65.25	64.29	65.95	57.54	72.01	0.6895
	Excl. significant	61.28	57.29	64.15	53.45	67.66	0.6330
	HRV	60.54	59.14	61.54	52.49	67.72	0.6308
MLP	All	75.61	68.79	80.51	71.71	78.26	0.8222
	Significant	60.03	68.79	53.74	51.68	70.57	0.6696
	Excl. significant	71.58	66.36	75.33	65.97	75.72	0.7769
	HRV	67.31	71.29	64.46	59.11	75.78	0.7364

Note: Four sets of input data were applied: all 33 multimodal features (All), six significant predictors from the linear regression analysis (Significant), 27 features excluding the six significant features (Excl. significant), and 24 HRV features excluding EDA, RESP, and temperature features (HRV).
Abbreviations: ACC, accuracy; SEN, sensitivity; SPE, specificity.

set excluded the six significant features and included the remaining 27 features. The final feature set only included the 24 HRV features and excluded the EDA, RESP, and temperature features.

The highest accuracy (75.61%) was achieved using the MLP with all 33 features; this algorithm also produced the highest specificity (80.51%), PPV (71.71%), NPV (78.26%), and AUC (0.8222). The highest sensitivity (71.50%) was achieved using the LoR algorithm with the significant features as input data. The LoR, KNN, SVM, and RF models exhibited the highest accuracy with the significant feature set; these results were better than those of the dataset without the significant features and the dataset with HRV features only.

We further investigated the effect of different combinations of input features on the performance of the classifiers. For this, we ranked the features based on MI. Table 6 shows the top 12 features ranked by the MI method, including 10 HRV features, 1 PT feature (rank 6), and 1 RESP feature (rank 12).

Next, we evaluated the performance of the LoR, KNN, and SVM classifiers using the Top 12, 9, 6, and 3 features as input data (Figure 4). The MLP and RF models were not tested with these feature sets because their performance was not substantially improved by reducing the number of features. MLP showed the highest accuracy (75.61%) with all 33 features, whereas RF showed 64.15% accuracy under the same conditions. The accuracy of RF was similar to its highest accuracy, 65.25%, which was achieved with the significant features. The line graphs in Figure 4 show the prediction accuracy of the LoR, KNN, and SVM classifiers with the top-ranked features according to the MI method. These accuracies were compared with those evaluated with all 33 features and the significant features for each algorithm

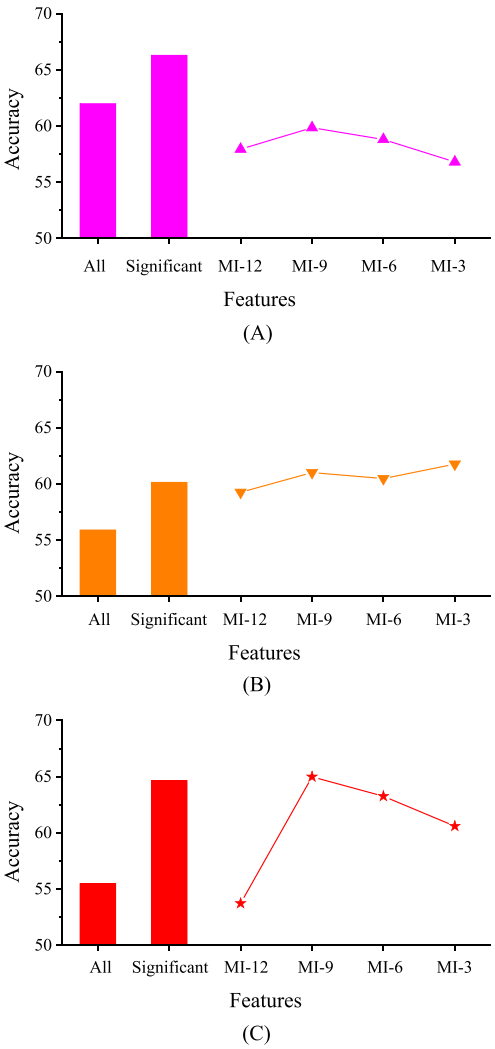

FIGURE 4 Classification accuracy for each combination of input dataset and machine learning algorithm. MI-12, MI-9, MI-6, and MI-3 represent the Top 12, 9, 6, and 3 features ranked using the MI method, respectively: (A) LoR, (B) KNN, (C) SVM

TABLE 6 Top 12 features ranked based on MI

Rank	Feature	MI
1	res_pNN50	0.09674
2	rec_RMSSD	0.08821
3	rec_SDNN	0.08818
4	rec_HF	0.08646
5	str_RMSSD	0.07791
6	rec_FT	0.06657
7	rec_pNN50	0.06572
8	res_LF	0.06486
9	str_VLF	0.06134
10	str_pNN50	0.05523
11	res_HF	0.05269
12	rec_RR	0.05070

(bar graphs in Figure 4). In LoR, the highest accuracy was 59.85%, which was achieved with the top nine features (MI-9); this accuracy was lower than that achieved with all features. In KNN, the accuracy increased to 61.76% as the number of features was reduced to three; this was slightly greater than the accuracy achieved with the significant features. In SVM, the highest accuracy was 64.99% with the MI-9 features, which was similar to the accuracy achieved with the significant features.

4 | DISCUSSION

We demonstrated that using machine learning models, patients with PD can be distinguished from HCs with an accuracy of 75.61% on the basis of physiological responses

induced by PPP tasks. The use of all 33 features extracted from four different physiological signals resulted in the highest accuracy in the classification of the PD and HC groups using MLP. Linear regression analysis showed that res_LF, res_FT, str_FT, rec_HR, str_HR, and rec_pNN50 were significant predictors of PD. These six significant features selected using linear regression were multimodal and included four HRV and two temperature indices. LoR, KNN, SVM, and RF exhibited the highest accuracy with the significant feature set, performing better than with the dataset excluding significant features and the dataset with HRV features only. These results suggest that combining multimodal physiological signals can improve the performance of classifiers and that the significant features found using linear regression analysis play important roles in classification.

Four HRV features—rec_pNN50, res_LF, rec_HR, and str_HR—were included in the six significant predictors of PD identified by the linear regression analysis. Abnormal autonomic activity represented by changes in HRV features has been studied in various anxiety disorders, including PD, GAD, posttraumatic stress disorder, and social anxiety disorder. Chalmers and others [14] performed a meta-analysis and found that PD was associated with reduced time-domain HRV features and HF. Previous studies have reported a significant reduction in pNN50 in patients with PD compared with HCs [18, 49, 50]. Consistently, PD was negatively related to pNN50 during the recovery phase in this study. pNN50 is a time-domain feature and reflects the activity level of the parasympathetic tone [51]. It is closely associated with HF [52]. In this study, PD was positively related to LF during the rest phase. Previous studies have shown that LF is significantly higher in patients with PD than in HCs [17, 50, 53]. LF is affected by both vagal and sympathetic activities [54], and an increased LF indicates elevated sympathetic activation. Therefore, our results suggest that PD is associated with decreased vagal tone and increased sympathetic activity.

Notably, rec_HR was positively related to PD, whereas str_HR was negatively related to PD. Cardiac responses during the stress and recovery phases can reflect altered autonomic activity in patients with PD. The generalized unsafety theory of stress (GUTS) provides an explanation of how abnormal cardiac activity and anxiety disorders might be related and proposes that the stress response is chronically inhibited while safety is perceived [55]. However, when a stressor is perceived, this cognitive inhibition is removed, and the default stress response is triggered, resulting in the activation of physiological responses, such as an increase in HR and a decrease in HRV. When the stressor is removed, safety is detected, the stress response is inhibited, and physiological activation returns to normal levels. However, chronically anxious individuals, such as patients with PD, have difficulty in perceiving safety, and, as a result, their stress response is not inhibited but remains active [5]. Therefore, anxious individuals may not fully recover from the stress response but remain chronically stressed [10]. GUTS differentiates between the stress response during exposure to a stressor and prolonged stress response, which occurs in the recovery phase after the stressor has been removed [56]. These results are consistent with the positive association observed between rec_HR and PD, indicating that individuals in the PD group might have experienced a prolonged stress response during the recovery phase even after the stress task had ended. However, the relationship between str_HR and PD was negative. This result was not consistent with the findings of previous studies, which reported significantly increased HR in patients with PD compared with HCs [25, 49]. Although a significant relationship between HR and PD during the stress phase was observed in the opposite direction, this finding might also reflect altered autonomic reactivity in patients with PD. In the current setup for experimental stress and recovery, the immediate stress response and prolonged stress response may lead to mixed outcomes in patients with PD. These results suggest the need for future studies on stress responses in patients with PD.

Two temperature indices were related to PD: res_FT was negatively related and str_FT was positively related to PD. FT is considered to be a surrogate marker of blood flow changes resulting from vascular reactivity and is influenced primarily by sympathetic adrenergic vasoconstrictor nerves [57]. With more tense muscles under strain, blood vessels contract, and FT decreases. For example, FT is significantly decreased by stresses such as emotional stress and fear and is increased during relaxation, boredom, and sleep [57–60]. However, the relationship between FT and PD is not clear. Freedman and others reported that patients with PD had significantly lower FT than healthy subjects [25]. In contrast, Pruneti and others found that FT was higher in patients with PD than in HCs in all three phases of the PPP test: rest, stress, and recovery [30]. We also found that res_FT and str_FT were related to PD, albeit oppositely. Pruneti and others also demonstrated that compared with the rest phase, the stress phase resulted in a decrease in FT in the HC group but an increase in FT in the PD group, suggesting that physiological responses to stress differed between patients with PD and healthy subjects [30]. An increase in FT during the stress phase in the PD group may be associated with elevated or dysfunctional vascular reactivity, which is considered to be one of the physiological factors affecting altered HRV in patients with PD.

The SCL and RESP features were not identified as significant predictors in the linear regression analysis. However, altered SCL in patients with PD has been considered an important physiological index related to increased tension and anxiety [61–63]. Pruneti and others showed that SCL was higher in patients with PD than in HCs in the rest, stress, and recovery phases and suggested that increased SCL can indicate an acute state of anxiety [30]. Patients with PD also exhibited increased respiratory irregularity and mean minute ventilation compared with HCs, suggesting a link between panic and respiratory abnormalities [64]. Nonetheless, in some previous studies, patients with PD showed no significant changes in SCL and respiratory features at rest compared with HCs [65, 66], suggesting that methodological differences can lead to different conclusions.

In this study, we only used SCL as the representative feature of the EDA signals. However, other EDA measures, such as phasic components, have also been used to evaluate autonomic activity. The lack of inclusion of these measures is a limitation of this study. To determine which feature should be extracted from the EDA signals, we took into account the extent to which the features had been studied with respect to autonomic activity in patients with PD. We selected SCL as the representative feature for EDA based on a previous study, which used an experimental design similar to ours [29]. Using a PPP method, Pruneti and others tested four patient groups; the patients in each group had either PD, GAD, major depression episode (MSE), or obsessive–compulsive disorder (OCD). In their study, SCL was high in the PD and GAD groups, whereas the MSE and OCD groups exhibited flat and nonreactive activation to stressful stimuli. Other studies have also demonstrated high SCL in patients with PD [67–71].

The use of EDA features based on phasic component analysis, such as skin conductance response (SCR), may improve the performance of classifier models. Hoehn-Saric and others [72] suggested that the SCR of patients with PD was greater and more variable than that of HCs during non-panic-inducing psychological stress. SCR is a more phasic measure of change in skin conductance and is related to the number of sweat glands that are activated in response to particular stimuli [73]. It has been used to evaluate subjects' responses to event-related experiments ("startle-like" stimuli) or tonic stimuli tests (a change in condition, workload, or cognitive stress) [74]. The results of these studies suggest that SCR can be used to investigate sweat responses to the stimulus that we used in our PPP experiment. In the future, we plan to test additional EDA features to conduct more in-depth comparisons between multimodal physiological features and to improve prediction performance.

The MLP using all 33 features outperformed the one using the HRV-only dataset (24 features), demonstrating that combining multimodal features improved the performance of MLP classifiers in this study. In addition, the use of the significant feature set resulted in the highest accuracy for LoR, KNN, SVM, and RF. We further tested the effects of different combinations of input features in LoR, KNN, and SVM using feature ranking based on MI estimation. Five HRV features and one temperature feature were selected as the top six ranked features (MI-6). However, there were no features common between the MI-6 and significant feature sets. In LoR, feature selection did not improve performance, and in KNN and SVM, feature selection did not result in substantially improved accuracy compared with the highest performance achieved with the significant feature set. These results suggest that the significant features can play an important role in detecting abnormal ANS activity in patients with PD.

MLP showed the highest accuracy in differentiating between the PD and HC groups. The classification accuracy depends not only on the algorithms but also on the nature of the datasets. Because the characteristics of the data, such as linearity and homogeneity, were unpredictable, we implemented and compared five machine learning algorithms having very distinct basic principles. Each algorithm has its own advantages and disadvantages [41, 75]. Our results do not suggest that MLP outperforms other algorithms when using physiological data obtained from different subjects or experimental setups. Other researchers have achieved good results with different algorithms. For instance, Na and others used LoR to achieve the best accuracy for discriminating PD from other anxiety disorders [31].

In the current study, the highest accuracy for classifying the individuals into PD and HC groups was 75.61%, which was achieved using MLP and multimodal physiological features. Na and others achieved 78.4% accuracy for distinguishing PD from other anxiety disorders using LoR and HRV data [31]. Lueken and others achieved 79% accuracy for distinguishing depressive comorbidity from PD using a tree ensemble classifier and fMRI data [32]. Sundermann and others achieved <60% accuracy for predicting treatment outcomes in patients with PD and agoraphobia using SVM and fMRI data [33]. However, because the previous studies were based on different types of data and did not focus on the differentiation between PD and HC groups, the comparisons of accuracy are complicated. We used nested CV to test the generalization performance. The nested CV was repeated 10 times to further improve the estimation. Therefore, our validation method did not overestimate the performance of the model. Hence, our results should be comparable to previously reported

automated discrimination of PD based on physiological or image data. Data obtained during PPP tasks can be considered sequential data with three time steps. In future studies, we will implement deep learning algorithms for sequence processing, with the expectation that accounting for temporal relationships between data points would improve the performance of classifiers [76, 77].

5 | CONCLUSIONS

We demonstrated that patients with PD can be differentiated from HCs with 75.61% accuracy using physiological features. We found that combining multimodal signals was crucial for identifying abnormal ANS reactivity in patients with PD and improving the performance of the classifiers. The features obtained from the stress and recovery phases were included as significant predictors. These findings suggest that multimodal physiological features measured during various states of the ANS have the potential to objectively differentiate patients with PD.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

ORCID

Eun Hye Jang https://orcid.org/0000-0001-8348-2663
Sangwon Byun https://orcid.org/0000-0002-5094-8372

REFERENCES

1. A. Y. Kim, E. H. Jang, S. Kim, K. W. Choi, H. J. Jeon, H. Y. Yu, and S. Byun, *Automatic detection of major depressive disorder using electrodermal activity*, Sci. Rep. **8** (2018), 17030.
2. S. Byun, A. Y. Kim, E. H. Jang, S. Kim, K. W. Choi, H. Y. Yu, and H. J. Jeon, *Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol*, Comput. Biol. Med. **112** (2019), 103381.
3. S. Byun, A. Y. Kim, E. H. Jang, S. Kim, K. W. Choi, H. Y. Yu, and H. J. Jeon, *Entropy analysis of heart rate variability and its application to recognize major depressive disorder: A pilot study*, Technol. Health Care **27** (2019), 407–424.
4. G. Valenza, L. Citi, C. Gentili, A. Lanatá, E. P. Scilingo, and R. Barbieri, *Point-process nonlinear autonomic assessment of depressive states in bipolar patients*, Methods Inf. Med. **53** (2014), 296–302.
5. P. de Jonge, A. M. Roest, C. C. Lim, S. E. Florescu, E. J. Bromet, D. J. Stein, M. Harris, V. Nakov, J. M. Caldas-de-Almeida, D. Levinson, A. O. al-Hamzawi, J. M. Haro, M. C. Viana, G. Borges, S. O'Neill, G. de Girolamo, K. Demyttenaere, O. Gureje, N. Iwata, S. Lee, C. Hu, A. Karam, J. Moskalewicz, V. Kovess-Masfety, F. Navarro-Mateu, M. O. Browne, M. Piazza, J. Posada-Villa, Y. Torres, M. ten Have, R. C. Kessler, and K. M. Scott, *Cross-national epidemiology of panic disorder and panic attacks in the world mental health surveys*, Depress. Anxiety **33** (2016), 1155–1177.
6. J. C. Ballenger, *Toward an integrated model of panic disorder*, Am. J. Orthopsychiatry **59** (1989), 284–293.
7. D. F. Klein and M. Fink, *Psychiatric reaction patterns to imipramine*, Am. J. Psychiatry **119** (1962), 432–438.
8. S. Freud, *The aetiology of hysteria*, In *The standard edition of the complete psychological works of Sigmund Freud*, J. Strachey (ed.), Hogarth Press, London, 1986, 162–222.
9. American Psychiatric Association, *Diagnostic and statistical manual for mental disorders*, 4th ed. DSM-IV, American Psychiatric Press, Washington DC, USA, 1994.
10. F. N. Busch, B. L. Milrod, and M. B. Singer, *Theory and technique in psychodynamic treatment of panic disorder*, J. Psychother. Pract. Res. **8** (1999), 234–242.
11. M. K. Hasan and R. P. Mooney, *Panic disorder: A review*, Compr. Ther. **12** (1986), 3–7.
12. R. R. Freedman, P. Ianni, E. Ettedgui, and N. Puthezhath, *Ambulatory monitoring of panic disorder*, Arch. Gen. Psychiatry **42** (1985), 244–248.
13. M. R. Liebowitz, J. M. Gorman, A. J. Fyer, M. Levitt, D. Dillon, G. Levy, I. L. Appleby, S. Anderson, M. Palij, S. O. Davies, and D. F. Klein, *Lactate provocation of panic attacks. II. Biochemical and physiological findings*, Arch. Gen. Psychiatry **42** (1985), 709–719.
14. J. A. Chalmers, D. S. Quintana, M. J. A. Abbott, and A. H. Kemp, *Anxiety disorders are associated with reduced heart rate variability: A meta-analysis*, Front. Psych. **5** (2014), 80.
15. P. J. Tully, S. M. Cosh, and B. T. Baune, *A review of the affects of worry and generalized anxiety disorder upon cardiovascular health and coronary heart disease*, Psychol. Health Med. **18** (2013), 627–644.
16. A. Pittig, J. J. Arch, C. W. R. Lam, and M. G. Craske, *Heart rate and heart rate variability in panic, social anxiety, obsessive-compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation*, Int. J. Psychophysiol. **87** (2013), 19–27.
17. H. Cohen, J. Benjamin, A. B. Geva, M. A. Matar, Z. Kaplan, and M. Kotler, *Autonomic dysregulation in panic disorder and in post-traumatic stress disorder: Application of power spectrum analysis of heart rate variability at rest and in response to recollection of trauma or panic attacks*, Psychiatry Res. **96** (2000), 1–13.
18. A. Garakani, J. M. Martinez, C. J. Aaronson, A. Voustianiouk, H. Kaufmann, and J. M. Gorman, *Effect of medication and psychotherapy on heart rate variability in panic disorder*, Depress. Anxiety **26** (2009), 251–258.
19. S. J. Petruzzello, D. M. Landers, B. D. Hatfield, K. A. Kubitz, and W. Salazar, *A meta-analysis on the anxiety-reducing effects of acute and chronic exercise. Outcomes and mechanisms*, Sports Med. **11** (1991), 143–182.
20. D. J. McEntee and R. P. Halgin, *Cognitive group therapy and aerobic exercise in the treatment of anxiety*, J. Coll. Stud. Psychother. **13** (1999), 37–55.
21. K. L. Rennie, H. Hemingway, M. Kumari, E. Brunner, M. Malik, and M. Marmot, *Effects of moderate and vigorous physical activity on heart rate variability in a British study of civil servants*, Am. J. Epidemiol. **158** (2003), 135–143.
22. M. E. Alvarenga, J. C. Richards, G. Lambert, and M. D. Esler, *Psychophysiological mechanisms in panic disorder: A correlative analysis of noradrenaline spillover, neuronal noradrenaline*

reuptake, power spectral analysis of heart rate variability, and psychological variables, Psychosom. Med. **68** (2006), 8–16.

23. A. Kotianova, M. Kotian, M. Slepecky, M. Chupacova, J. Prasko, and I. Tonhajzerova, *The differences between patients with panic disorder and healthy controls in psychophysiological stress profile*, Neuropsychiatr. Dis. Treat. **14** (2018), 435–441.
24. R. Vetrugno, R. Liguori, P. Cortelli, and P. Montagna, *Sympathetic skin response: Basic mechanisms and clinical applications*, Clin. Auton. Res. **13** (2003), 256–270.
25. R. R. Freedman, P. Ianni, E. Ettedgui, R. Pohl, and J. M. Rainey, *Psychophysiological factors in panic disorder*, Psychopathology **17** (1984), 66–73.
26. D. McDuff, S. Gontarek, and R. Picard, *Remote measurement of cognitive stress via heart rate variability*, (Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA), Aug. 2014, pp. 2957–2960.
27. G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, *Review on psychological stress detection using biosignals*, IEEE Trans. Affect. Comput. **13** (2019), 440–460.
28. P. Grossman, *Respiration, stress, and cardiovascular function*, Psychophysiology **20** (1983), 284–300.
29. C. Pruneti, C. Cosentino, M. Sgromo, and A. Innocenti, *Skin conductance response as a decisive variable in individuals with a DSM-IV TR axis I diagnosis*, JMED Res. **2014** (2014), 565009.
30. C. Pruneti, M. Saccò, C. Cosentino, and D. Sgromo, *Relevance of autonomic arousal in the stress response in psychopathology*, J. Basic Appl. Sci. **12** (2016), 176–184.
31. K. S. Na, S. E. Cho, and S. J. Cho, *Machine learning-based discrimination of panic disorder from other anxiety disorders*, J. Affect. Disord. **278** (2021), 1–4.
32. U. Lueken, B. Straube, Y. Yang, T. Hahn, K. Beesdo-Baum, H. U. Wittchen, C. Konrad, A. Ströhle, A. Wittmann, A. L. Gerlach, B. Pfleiderer, V. Arolt, and T. Kircher, *Separating depressive comorbidity from panic disorder: A combined functional magnetic resonance imaging and machine learning approach*, J. Affect. Disord. **184** (2015), 182–192.
33. B. Sundermann, J. Bode, U. Lueken, D. Westphal, A. L. Gerlach, B. Straube, H. U. Wittchen, A. Ströhle, A. Wittmann, C. Konrad, T. Kircher, V. Arolt, and B. Pfleiderer, *Support vector machine analysis of functional magnetic resonance imaging of interoception does not reliably predict individual outcomes of cognitive behavioral therapy in panic disorder with agoraphobia*, Front. Psych. **8** (2017), 1–11.
34. G. D. Fuller, *Biofeedback methods and procedures in clinical practice*, Biofeedback Press, San Francisco, USA, 1979.
35. T. Hoehn, S. Braune, G. Scheibe, and M. Albus, *Physiological, biochemical and subjective parameters in anxiety patients with panic disorder during stress exposure as compared with healthy controls*, Eur. Arch. Psychiatry Clin. Neurosci. **247** (1997), 264–274.
36. M. Lehofer, M. Moser, R. Hoehn-Saric, D. McLeod, P. Liebmann, B. Drnovsek, S. Egner, G. Hildebrandt, and H. G. Zapotoczky, *Major depression and cardiac autonomic control*, Biol. Psychiatry **42** (1997), 914–919.
37. P. Zarjam, J. Epps, F. Chen, and N. H. Lovell, *Estimating cognitive workload using wavelet entropy-based features during an arithmetic task*, Comput. Biol. Med. **43** (2013), 2186–2195.
38. M. K. Shear, T. A. Brown, D. H. Barlow, R. Money, D. E. Sholomskas, S. W. Woods, J. M. Gorman, and L. A. Papp, *Multicenter collaborative panic disorder severity scale*, Am. J. Psychiatry **154** (1997), 1571–1575.
39. M. Hamilton, *The assessment of anxiety states by rating*, Br. J. Med. Psychol. **32** (1959), 50–55.
40. J. F. Thayer, S. S. Yamamoto, and J. F. Brosschot, *The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors*, Int. J. Cardiol. **141** (2010), 122–131.
41. M. Liu, M. Wang, J. Wang, and D. Li, *Comparison of random forest, support vector machine and back propagation neural network for electronic tongue data classification: Application to the recognition of orange beverage and Chinese vinegar*, Sens. Actuators B **177** (2013), 970–980.
42. G. Sun, T. Shinba, T. Kirimoto, and T. Matsui, *An objective screening method for major depressive disorder using logistic regression analysis of heart rate variability data obtained in a mental task paradigm*, Front. Psych. **7** (2016), 180.
43. G. C. Cawley and N. L. C. Talbot, *On over-fitting in model selection and subsequent selection bias in performance evaluation*, J. Mach. Learn. Res. **11** (2010), 2079–2107.
44. 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 **6** (2017), 1–9.
45. A. M. Molinaro, R. Simon, and R. M. Pfeiffer, *Prediction error estimation: A comparison of resampling methods*, Bioinformatics **21** (2005), 3301–3307.
46. G. Chandrashekar and F. Sahin, *A survey on feature selection methods*, Comput. Electr. Eng. **40** (2014), 16–28.
47. A. Kraskov, H. Stögbauer, and P. Grassberger, *Estimating mutual information*, Phys. Rev. E - Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top. **69** (2004), 066138.
48. B. C. Ross, *Mutual information between discrete and continuous data sets*, PLoS ONE **9** (2014), e87357.
49. K. W. Choi, E. H. Jang, A. Y. Kim, M. Fava, D. Mischoulon, G. I. Papakostas, D. J. Kim, K. Kim, H. Y. Yu, and H. J. Jeon, *Heart rate variability for treatment response between patients with major depressive disorder versus panic disorder: A 12-week follow-up study*, J. Affect. Disord. **246** (2019), 157–165.
50. J. M. Martinez, A. Garakani, H. Kaufmann, C. J. Aaronson, and J. M. Gorman, *Heart rate and blood pressure changes during autonomic nervous system challenge in panic disorder patients*, Psychosom. Med. **72** (2010), 442–449.
51. G. G. Berntson, J. Thomas Bigger JR., D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. van der Molen, *Heart rate variability: Origins, methods, and interpretive caveats*, Psychophysiology **34** (1997), 623–648.
52. M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, and P. J. Schwartz, *Heart rate variability: Standards of measurement, physiological interpretation, and clinical use*, Eur. Heart J. **17** (1996), 354–381.
53. E.-H. Kang, I. S. Lee, J. E. Park, K. J. Kim, and B. H. Yu, *Platelet serotonin transporter function and heart rate variability in patients with panic disorder*, J. Korean Med. Sci. **25** (2010), 613–618.
54. J. M. Gorman and R. P. Sloan, *Heart rate variability in depressive and anxiety disorders*, Am. Heart J. **140** (2000), 77–83.

55. J. F. Brosschot, B. Verkuil, and J. F. Thayer, *Exposed to events that never happen: Generalized unsafety, the default stress response, and prolonged autonomic activity*, Neurosci. Biobehav. Rev. **74** (2017), 287–296.
56. J. F. Brosschot, *Markers of chronic stress: Prolonged physiological activation and (un)conscious perseverative cognition*, Neurosci. Biobehav. Rev. **35** (2010), 46–50.
57. E.-H. Jang, B. J. Park, M. S. Park, S. H. Kim, and J. H. Sohn, *Analysis of physiological signals for recognition of boredom, pain, and surprise emotions*, J. Physiol. Anthropol. **34** (2015), 25.
58. U. Miura, *The effect of variations in relative humidity upon skin temperature and sense of comfort*, Am. J. Epidemiol. **13** (1931), 432–459.
59. F. B. Talbot, V. Bates, E. Bates, and A. J. Dalrymple, *Skin temperatures of children*, Am. J. dis. Child. **42** (1931), 965–967.
60. H. Helson and L. Quantius, *Changes in skin temperature following intense stimulation*, J. Exp. Psychol. **17** (1934), 20–35.
61. J. A. Gray, *The neuropsychology of anxiety*, Br. J. Psychol. **69** (1978), 417–434.
62. D. C. Fowles, *The three arousal model: Implications of Gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy*, Psychophysiology **17** (1980), 87–104.
63. D. C. Fowles, *Psychophysiology and psychopathology: A motivational approach*, Psychophysiology **25** (1988), 373–391.
64. M. Grassi, D. Caldirola, G. Vanni, G. Guerriero, M. Piccinni, A. Valchera, and G. Perna, *Baseline respiratory parameters in panic disorder: A meta-analysis*, J. Affect. Disord. **146** (2013), 158–173.
65. R. Hoehn-Saric and D. R. McLeod, *Somatic manifestations of normal and pathological anxiety*, In *Biology of anxiety disorders*, R. Hoehn-Saric, D. R. McLeod (eds.), American Psychiatric Association, Arlington, VA, USA, 1993, 177–222.
66. M. B. Stein and G. J. C. Asmundson, *Autonomic function in panic disorder: Cardiorespiratory and plasma catecholamine responsivity to multiple challenges of the autonomic nervous system*, Biol. Psychiatry **36** (1994), 548–558.
67. W. T. Roth, M. J. Telch, C. B. Taylor, J. A. Sachitano, C. C. Gallen, M. L. Kopell, K. L. McClenahan, W. S. Agras, and A. Pfefferbaum, *Autonomic characteristics of agoraphobia with panic attacks*, Biol. Psychiatry **21** (1986), 1133–1154.
68. W. T. Roth, A. Ehlers, C. B. Taylor, J. Margraf, and W. S. Agras, *Skin conductance habituation in panic disorder patients*, Biol. Psychiatry **27** (1990), 1231–1243.
69. L. Dratcu and A. Bond, *Panic patients in the non-panic state: Physiological and cognitive dysfunction*, Eur. Psychiatry **13** (1998), 18–25.
70. A. C. B. V. Parente, C. Garcia-Leal, C. M. del-Ben, F. S. Guimarães, and F. G. Graeff, *Subjective and neurovegetative changes in healthy volunteers and panic patients performing simulated public speaking*, Eur. Neuropsychopharmacol. **15** (2005), 663–671.
71. C. Pruneti, F. Fontana, and C. Fante, *Autonomic changes and stress response in psychopathology*, Int. J. Psychophysiol. **69** (2008), 224–225.
72. R. Hoehn-Saric, D. R. McLeod, and W. D. Zimmerli, *Psychophysiological response patterns in panic disorder*, Acta Psychiatr. Scand. **83** (1991), 4–11.
73. W. Boucsein, *Electrodermal activity*, Springer, NY, USA, 2012.
74. H. F. Posada-Quintero and K. H. Chon, *Innovations in electrodermal activity data collection and signal processing: A systematic review*, Sensors. **20** (2020), 479.
75. S. Uddin, A. Khan, M. E. Hossain, and M. A. Moni, *Comparing different supervised machine learning algorithms for disease prediction*, BMC Med. Inform. Decis. Mak. **19** (2019), 1–16.
76. A. Onan and M. A. Tocoglu, *A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification*, IEEE Access **9** (2021), 7701–7722.
77. A. Onan, *Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks*, Concurr. Comput. Pract. Exp. **33** (2021), 1–12.

AUTHOR BIOGRAPHIES

Eun Hye Jang received her BA, MA, and PhD degrees in experimental and biological psychology from Chungnam National University, Daejeon, South Korea, in 2000, 2002, and 2009, respectively. Since 2009, she has worked as a senior researcher at the Electronics and Telecommunications Research Institute. Her research interests include the recognition of emotional and psychiatric states based on psychophysiology and ICT-cognition/emotion convergence.

Kwan Woo Choi received his BS degree in medicine from the School of Medicine, Sungkyunkwan University, Seoul, South Korea, in 2007 and his MS and PhD degrees in psychiatry from the School of Medicine, Sungkyunkwan University, Seoul, South Korea, in 2012 and 2020, respectively. From 2017 to 2019, he worked for Samsung Medical Center, Seoul, South Korea. From 2019 to 2022, he was affiliated with the Department of Psychiatry, Korea University, Seoul, South Korea, as a clinical assistant professor. His main clinical and research interests are depression, anxiety disorders, and suicide.

Ah Young Kim received her BS and MS degrees in electrical engineering from Soongsil University, Seoul, South Korea, in 2006 and 2012, respectively. Since 2012, she has been affiliated with the Electronics and Telecommunications Research Institute, Daejeon, South Korea, where she is currently a researcher in the Medical Sensor Research Section. Her research interests include signal processing of emotion recognition using bio-signals.

Han Young Yu received his BS degree in physics from the Department of Physics, Yonsei University, Seoul, South Korea, in 1997 and his MS and PhD degrees in physics from the Department of Physics, Seoul National University, Seoul, South Korea, in 1999 and 2003, respectively. Since 2003, he has worked for the Electronics and Telecommunications Research Institute, Daejeon, South Korea, where he is a principal researcher. His main research interests are biomedical and livestock engineering.

Hong Jin Jeon received his BS degree from the Seoul National University College of Medicine, Seoul, South Korea, in 1997 and his MS in the field of psychiatry from the Seoul National University College of Medicine in 2005 and PhD in 2007. Since 2008, he has worked in the Department of Psychiatry, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea, where he is now a professor. He is now the Vice Dean of Research at Sungkyunkwan University School of Medicine, Director of the Digital Therapeutics Research Center at Samsung Medical Center and professor at Samsung Advanced Institute for Health Sciences & Technology. His main research interests are depression and digital therapeutics.

Sangwon Byun received his BS and MS degrees in electrical engineering and computer science from Seoul National University, Seoul, South Korea, in 2002 and 2004, respectively, and his PhD degree in the same major from Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2010. He was formerly a postdoctoral associate (2010–2014) and research associate (2014–2015) at MIT's Department of Biological Engineering. From 2015 to 2017, he was affiliated with Electronics and Telecommunications Research Institute, Daejeon, South Korea, as a senior researcher. Since 2017, he has been affiliated with the Department of Electronics Engineering at Incheon National University, Incheon, South Korea, where he is an associate professor. His research interests include digital healthcare technologies.

How to cite this article: E. H. Jang, K. W. Choi, A. Y. Kim, H. Y. Yu, H. J. Jeon, and S. Byun, *Automated detection of panic disorder based on multimodal physiological signals using machine learning*, ETRI Journal **45** (2023), 105–118. https://doi.org/10.4218/etrij.2021-0299