Intelligent Computing
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

Exploring Electroencephalography-Based Affective Analysis and Detection of Parkinson's Disease

Ravikiran Parameshwara[1†], Soujanya Narayana[1†], Murugappan Murugappan[2,3,4], Ibrahim Radwan[1], Roland Goecke[1], and Ramanathan Subramanian[1*]

[1]Human-Centred Technology Research Cluster, University of Canberra, Canberra, Australia. [2]Department of Electronics and Communication Engineering, Kuwait College of Science and Technology, Kuwait City, Kuwait. [3]Department of Electronics and Communication Engineering, Vels Institute of Sciences, Technology, and Advanced Studies, Chennai, India. [4]Centre for Excellence in Unmanned Aerial Systems, Universiti Malaysia Perlis, Perlis, Malaysia.

*Address correspondence to: ram.subramanian@canberra.edu.au
†These authors contributed equally to this work.

Citation: Parameshwara R, Narayana S, Murugappan M, Radwan I, Goecke R, Subramanian R. Exploring Electroencephalography-Based Affective Analysis and Detection of Parkinson's Disease. *Intell. Comput.* 2024;3:Article 0084. https://doi.org/10.34133/icomputing.0084

Submitted 20 July 2023
Revised 12 December 2023
Accepted 8 February 2024
Published 17 October 2024

While Parkinson's disease (PD) is typically characterized by motor disorder, there is also evidence of diminished emotion perception in PD patients. This study examines the utility of electroencephalography (EEG) signals to understand emotional differences between PD and healthy controls (HCs), and for automated PD detection. Employing traditional machine learning and deep learning methods on multiple EEG descriptors, we explore (a) dimensional and categorical emotion recognition and (b) PD versus HC classification from multiple descriptors characterizing emotional EEG signals. Our results reveal that PD patients comprehend arousal better than valence and, among emotion categories, fear, disgust, and surprise less accurately, and sadness most accurately. Mislabeling analyses confirm confounds among opposite-valence emotions for PD data. Emotional EEG responses also achieve near-perfect PD versus HC recognition. Cumulatively, our study demonstrates that (a) examining implicit responses alone enables (i) discovery of valence-related impairments in PD patients and (ii) differentiation of PD from HC and that (b) emotional EEG analysis is an ecologically valid, effective, practical, and sustainable tool for PD diagnosis vis-à-vis self-reports, expert assessments, and resting-state analysis.

Introduction

Parkinson's disease (PD) is a neurodegenerative disorder of the central nervous system that affects movements, often causing tremors. PD is characterized by the progressive loss of dopaminergic neurons in the substantia nigra [1]. In addition to motor dysfunctions, cognitive, behavioral, and emotional defects are common in PD [2], affecting over 10 million people globally [3].

A number of studies have examined motor and cognitive impairments in PD patients by examining explicit user responses (e.g., performance in recognition tasks and self-reports) or implicit responses such as electroencephalography (EEG) signals [4]. Some works detect PD from abnormalities in resting-state EEG [5,6]. Resting-state EEG is acquired in a highly controlled setting, e.g., requiring the subject to remain motionless with eyes closed in a dim and quiet room, which makes this setting ecologically invalid. A more realistic setting involves EEG acquisition during routine tasks such as music listening [7] or movie watching [8].

As movie and musical stimuli are often emotion eliciting [9], they enable researchers to understand how PD patients perceive emotions. Apart from specific emotions, some studies focus on PD perception of the valence (feeling of pleasantness/aversion) and arousal (emotional intensity) dimensions [10]. Prior studies show that PD patients have a deficit in recognizing positive and negative valence emotions from prosody [11] and facial expressions [12], and reduced reactivity to highly arousing pictures [13]. Recognizing emotions is critical to social interaction and communication, apart from inferring nonverbal behavior such as emotional voice and facial expressions [14,15].

Implicit physiological and biological signals reflect the characteristic activity of the central nervous system and cannot be intentionally suppressed. Recent studies have extensively employed biosignals [16,17] for emotion perception in healthy subjects. EEG, functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and positron emission tomography (PET) provide more reliable information on emotional states compared to other modalities [18]. EEG is noninvasive, has high temporal resolution, and can detect changes in brain activity over a span of milliseconds. EEG frequency bands are known to correlate with emotions [19]. Hand-coded EEG descriptors such as spectral power vectors (SPVs) enable emotion

detection, while convolutional neural networks (CNNs) can automatically learn cognitive and emotional correlates [17,20].

This study examines EEG-based PD emotion perception via a comparative analysis of data acquired from PD patients vis-à-vis healthy controls (HCs). We explore both low-level EEG descriptors such as SPVs and common spatial patterns (CSPs), and the intermediate EEG image and movie representations [20] to this end. SPVs represent the power distribution across different frequency bands, while CSPs are popular descriptors for maximizing interclass discriminability [21]. EEG image descriptors are synthesized via topomaps capturing both spectral and spatial activity, whereas the EEG movie descriptor additionally captures temporal activity trends. We employ classical machine learning (ML) and deep learning frameworks such as 1D-, 2D-, and 3D-CNNs for emotion decoding. As shown in Fig. 1, we perform categorical and dimensional emotion recognition (binary valence and arousal classification) and PD versus HC recognition from EEG data.

The key findings from our study are as follows: (a) Dimensional analysis reveals that arousal is better perceived in PD than valence; similar or superior classification is achieved with HC data for both attributes. (b) Fine-grained analyses of emotion class mislabeling reveal confounds among opposite-valence emotions for PD data; this trend is not discernible for HC. (c) Near-ceiling PD versus HC classification ($F1 \geq 0.97$) is achieved with a 2D-CNN on emotional EEG data, implying that affective neural responses of PD and HC subjects are highly discriminable. Analyzing emotional neural responses can, therefore, enable facile PD diagnosis and treatment. Our study makes the following contributions:

- It examines (a) PD versus HC recognition and (b) emotion perception in PD exclusively from EEG classification trends. While resting-state EEG analysis has achieved high PD recognition accuracy [6,22], it requires EEG acquisition in a highly controlled setting. In contrast, we examine EEG signals acquired during the routine task of emotional media viewing, which additionally enables PD emotion understanding.
- PD diagnosis and treatment heavily rely on patient self-reports and expert assessments. While the importance of preclinical diagnosis and the need for objective monitoring with wearables [23] has been highlighted recently, high PD versus HC discriminability achieved with passive data acquisition during a routine task points to a promising alternative.
- We employ multiple (a) EEG descriptors and (b) machine and deep learning recognition frameworks for analyses. Among the EEG descriptors, CSPs and SPVs are optimal for emotion and PD recognition, respectively. CNN frameworks trained with intermediate EEG image and movie descriptors, however, achieve superior emotion and PD recognition performance.

To highlight the novelty of our study, we review related work on (a) emotional impairments in PD patients and (b) the use of biosignals to assess emotional impairments.

Emotional impairments in PD patients

PD patients show not only motor symptoms but also cognitive [24] and emotional [25] deficits. Kan et al. [26] report PD deficits in recognizing the fear and disgust facial emotions. Clark et al. [12] note impaired anger recognition in PD patients with left hemisphere pathology and reduced surprise recognition with right hemisphere pathology. Baggio et al. [27] observe PD deficits in recognizing sadness, anger, and disgust, while Narme et al. [28] note impaired recognition of anger and fear. A meta-analysis indicates an initial PD deficit for negative emotions [25] and later for positive emotions [29]. Some studies employ nonvisual stimuli, e.g., auditory and verbal, to assess PD emotion deficits. In an emotional voice test [30], PD patients in general exhibit impaired recognition and expression. Kan et al. [26] observe reduced recognition of fear, surprise, and disgust from text.

Using biosignals to assess impairments

Emotion is a psychophysiological expression related to external stimuli, mood, and personality [9,16]. Wearable sensing technologies can help examine biosignals and interpret associated emotions. Various physiological signals (or biosignals) such as EEG, MEG, fMRI, facial (micro)expressions, eye movements, electrodermal activity, and heart rate have been employed to study emotions induced by stimuli, such as facial imagery [31], audiovisual music and movie clips [9,16], and advertisements [17], or impairments thereof.

fMRI brain activations reveal stronger activation in somatosensory regions during emotion processing for PD patients [32]. fMRI analyses show reduced functional activity in the left and right posterior putamen [33], disturbing emotional processing. Spontaneous facial expressivity in PD observed via

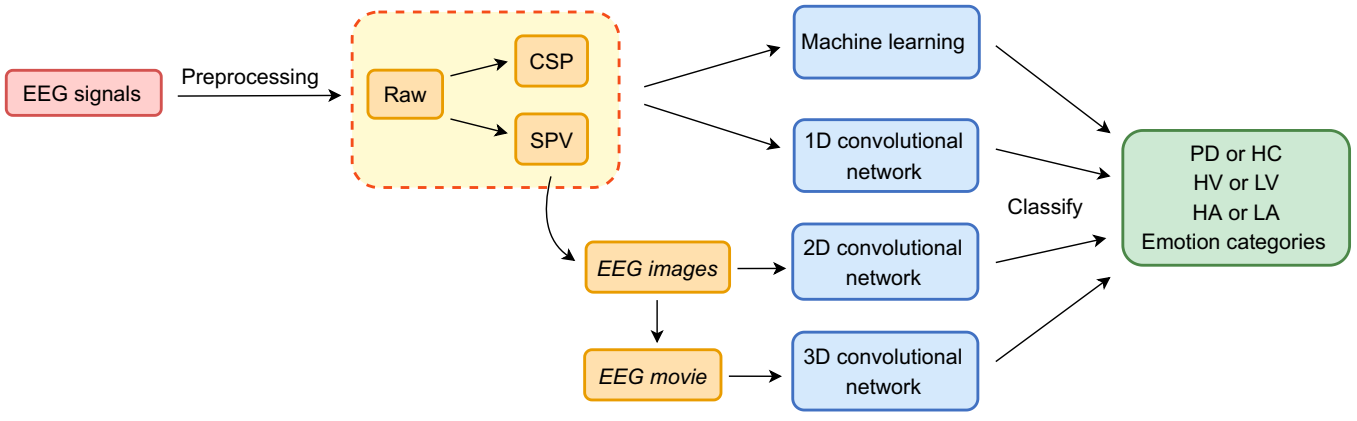

Fig. 1. Overview of our pipeline. EEG preprocessing and extraction of low-level features such as SPVs and CSPs is followed by feeding of these features or intermediate EEG image/movie representations to machine and deep learning frameworks to perform dimensional and discrete emotion recognition, and PD versus HC classification (HV and LV refer to high and low valence, while HA and LA refer to high and low arousal). Emotions belonging to each category are described in the "Valence classification" section.

electromyogram (EMG) and electrocardiogram (ECG) signals reveal differences between PD patients and controls [34]. A neuroimaging study involving fearful faces notes that dopamine levels modulate the amygdala's response in PD [35], and amygdala dysfunction induces impaired reactions to fear-inducing stimuli.

CNNs have been employed for EEG-based seizure prediction [36], fMRI-based schizophrenia detection [37], etc. Deep learning methods learn salient and latent neural representations [38]. EEG-based categorical emotion recognition in PD patients has been pursued via higher-order spectral statistics [8,39–41]. PD recognition via spectral analysis of resting-state EEG has been performed with k-nearest neighbor (kNN) and support vector machine (SVM) classifiers [22]. A 13-layer 1D-CNN for PD versus HC classification with resting-state EEG has been proposed in [6]. Binary PD versus HC EEG classification using a convolution recurrent neural network (CRNN) has been proposed in [42].

Identifying research gaps

While it is largely known that PD patients face emotional deficits, only a few studies examine these deficits via implicitly acquired biosignals, such as EEG. Moreover, prior studies on automated PD diagnosis have only examined resting-state EEG, but not emotional EEG signals. Emotional EEG signals can be (a) acquired via easy-to-use, portable headsets under routine settings and (b) utilized for PD recognition plus studying PD emotion deficits as described in this work. Table 1 compares and contrasts our work against the literature; evidently, our approach achieves optimal PD recognition with ecologically valid data, thus differs from prior work on resting-state and emotional EEG.

Materials and Methods

Dataset

The dataset comprises EEG signals from 20 nondemented PD (10 male/10 female) and 20 HC (9 male/11 female) subjects from Hospital Universiti Kebangsaan Malaysia, Kuala Lumpur, upon ethics approval (approval no. UKM1.5.3.5/244/FF-354-2012) [39,41]. The demographic and PD severity of the patients are included in Table S1. EEG data were recorded via the 14-channel wireless Emotiv Epoc headset (128-Hz sampling rate). Audiovisual stimuli were used to induce the 6 Ekman emotions (sadness, happiness, fear, disgust, surprise, and anger), resulting in a total of 1,440 samples (2 classes × 20 subjects × 6 emotions × 6 trials/emotion). The emotional stimuli were gathered from video clips collected from various sources on the internet, the International Affective Picture System (IAPS) [43] database, and the International Affective Digitized Sounds (IADS) [44] database. The IAPS, designed for studies on emotion, is a database comprising 956 pictures of everyday objects and scenes (e.g., household furniture). The IADS database comprises 935 digitally recorded sounds from daily life (e.g., footsteps and babies crying). The pictures and the sounds in the IAPS and IADS databases are annotated for valence and arousal, respectively. The stimuli employed for this study were synthesized from these 2 databases. A pair of stimuli with identical parity of valence and arousal was combined and presented to the participant. For example, a negative-high aroused sound was given along with a negative-high aroused image [40]. Each trial (stimulus episode) lasted 4 to 5 min. PD patients were optimally medicated to reduce tremors, and informed consent was obtained from all participants. Data acquisition, PD clinical history, ethics approval, and PD inclusion and exclusion criteria are described in [8,39].

Data preprocessing

EEG outlier samples were discarded by limiting the signal amplitude to 85 μV, followed by an infinite impulse response (IIR) bandpass Butterworth filter to retain the 8- to 49-Hz range [8]. We consider 8 to 49 Hz as the frequency range in this study as the α (8 to 13 Hz), β (13 to 30 Hz), and γ (30 to 49 Hz) spectral bands play an important role in emotion-related activities, as compared to the δ (1 to 4 Hz) and θ (4 to 8 Hz) bands [40]. Additionally, the β band activity (which increases in the basal ganglia and decreases in the motor cortex) is observed to contribute greatly to motor symptoms in PD [45].

Each filtered EEG signal, recorded for 40 to 50 s [8,39], was segmented into 5 s epochs to preserve temporal information following [46], resulting in 8 to 10 segments per recording. This process resulted in an average of 9.16 segments for 1,440 samples, leading to a total of 13,193 samples for feature extraction. Each of these samples, termed as raw data henceforth, are of dimension (14, 640). For classical ML methods, raw features were z-normalized followed by principal components analysis (PCA) to retain 95% data variance (PCA was not part of the CNN pipeline).

Feature extraction from raw EEG

Spectral power vector

Power spectral analysis was performed to estimate EEG spectral density upon spectral transformation [47]. On each epoch, a Butterworth bandpass filter was applied to extract the α (8 to 13 Hz), β (13 to 30 Hz), and γ (30 to 49 Hz) spectral bands. A fast Fourier transform (FFT) was performed, followed by summation of squared FFT values within each of the 3 frequency bands over the 14 electrodes to obtain the concatenated SPV $[\alpha_1, \beta_1, \gamma_1, ..., \alpha_{14}, \beta_{14}, \gamma_{14}]$. Thus, each raw EEG sample of dimension (14, 640) was transformed to (1, 42).

Common spatial patterns

CSPs were extracted by learning a linear combination of the original features [48]. Filters (transformations) were designed so that the transformed signal variance was maximal for one class and minimal for the other. Apart from dimensionality reduction, CSPs enable recovery of the original signal by gathering relevant information spread over multiple channels and are, hence, popular EEG features [47]. We learn the spatial transform w, which maximizes the function:

$$J_{CSP}(w) = \frac{wX_1X_1^Tw^T}{wX_2X_2^Tw^T} = \frac{wC_1w^T}{wC_2w^T}, \quad (1)$$

where C_i and X_i are, respectively, the spatial covariance matrix and the bandpass-filtered signal matrix for class i. In Eq. 1, wX_i is the spatially filtered EEG signal for class i and $wX_iX_i^Tw^T$ is the transformed signal variance, i.e., the band power of the filtered signal. Thus, maximizing $J_{CSP}(w)$ leads to spatially filtered signals whose interclass band power ratio is maximal, and can be solved via eigenvalue decomposition. The spatial filters w that

Table 1. Overview of studies examining PD behavior and automated PD detection

Behavior studied	References	Year	Description	Findings	Remarks
Implicit	[66]	1995	PD and HC subjects presented with emotional faces, and asked to pose those expressions.	PD patients were impaired relative to controls on making emotional faces.	Purely behavioral study, automated PD detection not attempted.
Explicit + Implicit	[62]	1996	PD and HC subjects shown emotional video clips and their emotional reactions encoded via Facial Action Coding System. Emotional self-ratings compiled.	PD group showed considerably less facial activity than controls.	Purely behavioral study, automated PD detection not attempted.
Implicit	[67]	2003	PD and HC subjects presented with odors and their emotional reactions encoded via Facial Action Coding System.	Spontaneous facial activity, ability to pose and mask facial expressions impaired in PD.	Purely behavioral study, automated PD detection not attempted.
Explicit + Implicit	[13]	2009	24 PD and 24 HC subjects presented with emotional pictures, and responses acquired via self-ratings (explicit) and EMG activity (implicit).	Reduced PD reactivity to low-valence, high-arousal pictures; behavior was not specific to any emotion category (e.g., fear and disgust).	Behavioral study with no automated PD detection.
Explicit	[11]	2010	PD and HC subjects presented with videos or text and asked to label the stimulus emotion. Emotion recognition compared via standardized tests across the lexical-semantic, prosody, and facial information channels.	PD emotion recognition capability increased with more channels, but PD group performed worse than controls across all channels.	Behavioral study with no automated PD detection.
Implicit (emotional EEG)	[58]	2014	PD versus HC recognition via emotional EEG analysis with SVM classification. Dataset proposed in [39] examined.	Mean accuracy of 87.9% achieved.	Routine and ecologically valid setting. Only PD versus HC classification attempted.
Implicit (rest-state EEG)	[6,22]	2019, 2020	PD versus HC recognition via resting-state EEG analysis using ML [22] and deep learning [6] approaches. Dataset proposed in [22] examined.	Mean recognition accuracy of 99.1% achieved with SVM [22], and 88.3% achieved with 1D-CNN [6].	Resting-state EEG compiled under highly controlled conditions, ecologically invalid.
Implicit (emotional EEG)	[68]	2022	Categorical emotion detection in PD and HC via emotional EEG analysis with a convolutional recurrent neural network. Dataset proposed in [39] and 2 other publicly available datasets examined.	Mean categorical emotion recognition accuracy of 83.2% achieved for PD data, and 86.0% achieved for HC data.	Routine and ecologically valid setting. Only emotion classification in PD and HC attempted.
Implicit (emotional EEG)	Our work		Dimensional and categorical emotion plus PD recognition via machine and deep learning approaches. Dataset in [39] examined.	Valence-specific mislabeling observed with PD data, while no arousal-related differences noted between PD and HC groups. Maximum accuracy/F1 scores of 93%, 98%, and 99% achieved for valence, arousal, and PD recognition, respectively.	PD and emotion EEG recognition with multiple classification methods. PD recognition higher accuracy with emotional EEG and comparable to prior work with resting-state EEG.

maximize $J_{CSP}(w)$ are the eigenvectors of the highest and lowest eigenvalues for matrices C_1, C_2. Hence, w gives feature vectors that are optimal for class discrimination. For each class i, the variances of only a small number of signals most suitable for distinguishing are used. Six filters corresponding to the 3 largest and smallest eigenvalues are used for generating the CSP feature of dimension (1, 6) for each (14, 640) EEG epoch, given by

$$f = \log\left(wC_2w^T\right) = \log\left(var\left(wX_2\right)\right). \quad (2)$$

It is shown that the classification accuracy does not improve with a larger number of filters [49].

Classical ML algorithms

Raw EEG data or extracted features were input to machine/deep learning classifiers (Fig. 1). We explore the following ML algorithms for classification.

- kNN, where the test sample is assigned the label corresponding to the mode of its k-closest neighbors based on a suitable distance metric [50].
- SVM, where input data are transformed to a high-dimensional space where the 2 classes are linearly separable and the interclass distance is maximal [51].
- Gaussian naive Bayes (GNB), a generative classifier assuming class-conditional feature independence [52].
- Decision tree (DT), which uses a tree-like graph structure where each leaf node represents a category label [53].
- Linear discriminant analysis (LDA), which linearly transforms data to achieve maximal interclass distance [54].
- Logistic regression (LR), which maps the input to class labels via the sigmoid function [55].

Model parameters are fine-tuned via grid search, upon performing 10-fold cross-validation (10FCV) on the training set. Table 2 presents the range of parameters explored for each ML algorithm.

Convolutional neural network pipeline

We explored 1D-, 2D-, and 3D-CNNs to learn EEG representations. Raw or extracted EEG features were fed to the 1D-CNN; the feature dimensions input to the 1D-CNN with raw, spectral, and CSP descriptors were (640, 14), 42, and 6, respectively. However, this representation ignores the EEG spatial structure; therefore, we synthesized the EEG image and EEG movie descriptors to preserve the spatial structure.

Extracted SPVs were transformed to an EEG image as in [20]. EEG electrodes distributed on the scalp in 3D were projected onto a 2D surface to capture the spatial activity distribution. Azimuthal equidistant projection was used to preserve the relative interelectrode distance. Scattered scalp power measurements were interpolated to derive a 32×32 pixel EEG image. Repeating this process for the α, β, and γ bands produced 3 topo-maps, which were then merged to form a 3-channel ($32 \times 32 \times 3$) EEG image [20]. To learn the temporal EEG structure, given that 3D-CNNs effectively learn from video chunks [56], we synthesized EEG movie samples comprising 5 images generated by sliding nonoverlapping 1 s windows over the 5 s epoch. The 3D-CNN input dimensionality is $5 \times 32 \times 32 \times 3$.

With the recent success of CNNs for PD diagnosis, we exploit an architecture inspired by [6] for 1D, 2D, and 3D data. The general architecture of the 3-layered 1D/2D/3D-CNN employed

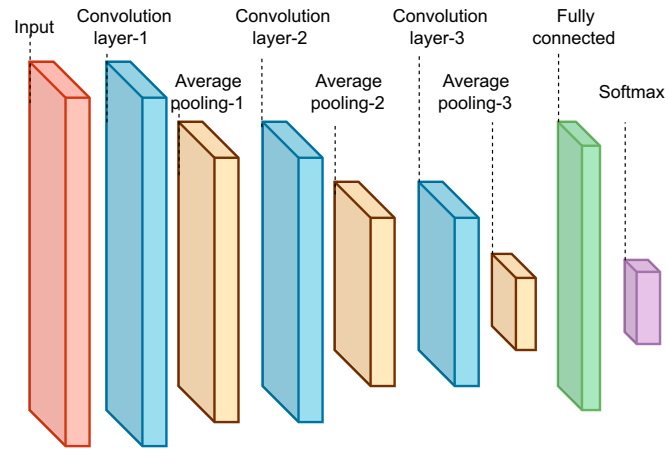

Fig. 2. Basic architecture of the 1D/2D/3D-CNN.

Table 2. Parameters and their ranges considered for grid search optimization

Algorithm	Parameters	Values
kNN	Number of neighbors	$2, 3, \cdots, 10$
	Distance metric	Euclidean, Chebychev, Minkowski
	Weight function	Uniform weight function, distance weight function
SVM	Regularization parameter (C)	$10^{-3}, 10^{-2}, \cdots, 10^{3}$
GNB	Variance smoothing factor	$10^{-3}, 10^{-2}, \cdots, 10^{3}$
DT	Maximum depth of the tree	Evenly spaced 50 numbers in the log-scale $[10^{-2}, 10]$
	Function to measure the quality of the split	Gini impurity, entropy
LDA	Solver	Singular value decomposition, least squares solution, eigenvalue decomposition
LR	Norm for the penalty	l_2, No penalty
	Inverse of the regularization strength (C)	$10^{-2}, 10^{-1}, \cdots, 10^{3}$

Table 3. Output dimensions of each layer in 1D-, 2D-, and 3D-CNN

	1D-CNN				
Layer	Spectral	CSP	Raw	2D-CNN	3D-CNN
Convolution layer-1	42, 16	6, 16	640, 16	32, 32, 16	5, 32, 32, 16
Average pooling-1	21, 16	3, 16	320, 16	16, 16, 16	3, 16, 16, 16
Convolution layer-2	21, 32	3, 32	320, 32	16, 16, 32	3, 16, 16, 32
Average pooling-2	11, 32	2, 32	160, 32	8, 8, 32	2, 8, 8, 32
Convolution layer-3	11, 32	2, 32	160, 32	8, 8, 32	2, 8, 8, 32
Average pooling-3	6, 32	1, 32	80, 32	4, 4, 32	1, 4, 4, 32
Flatten	192	32	2,560	512	512
Batch normalization	192	32	2,560	512	512
Fully connected	128	128	128	128	128
Softmax	2 or 6	2 or 6	2 or 6	2 or 6	2 or 6

for classification is shown in Fig. 2. Output dimensions for each CNN layer are presented in Table 3. Three convolutional layers convolve the input signal with a stride of 3 and comprise 16, 32, and 32 filters of size 3, 3×3, and $3 \times 3 \times 3$, respectively. Each convolutional layer is followed by average pooling over 2-unit regions. Batch normalization is applied to normalize prior activations, and a dropout of 0.1 to 0.5 is employed for regularization. The dense layer comprises 128 neurons and is followed by a softmax layer composed of 2 neurons (for dimensional emotion and PD versus HC classification) or 6 neurons (for categorical emotion recognition) conveying class probabilities. CNN hyperparameters (learning rate $\in[10^{-5}, \ldots, 10^{-1}]$, optimizer $\in${SGD, Adam, RMSPropogation}, and dropout rate) were tuned via 10FCV identical to the ML models.The code is available online at https://github.com/ravikiranrao/EEG-Parkinson.

Performance evaluation

For all results, we report the weighted F1 measure or the weighted mean of the per-class F1 scores, which accounts for the class imbalance noted in valence and arousal classification.

Results and Discussion

This section details classification results on PD and HC data for the valence and arousal dimensions [10] (val_results, asl_results), followed by categorical emotion (multi_results), and PD versus HC recognition (pdnc_results).

Valence classification

To examine emotional perception in PD, we first perform valence classification by training binary classifiers with (a) PD, (b) HC, and (c) combined PD and HC, or full EEG data. Happiness and surprise are categorized as high valence (HV), while sadness, fear, disgust, and anger data are grouped in the low valence (LV) category. The HV:LV class ratio within PD, HC, and the full data is 1:2.

Valence classification results on PD, HC, and the full data for various models are presented in Table 4. Higher F1 scores were achieved with HC data, implying reduced discriminability with PD EEG data. We discuss the results below. The valence classification results of all ML algorithms on PD, HC, and full data are shown in Table S2.

Classification with PD data

The impact of descriptors on the efficacy of ML techniques is evident from Table 4. A one-way analysis of variance (ANOVA) to examine the effect of features (Raw, SPV, and CSP) on F1 scores confirms the impact of descriptor type ($F(2, 27) = 969.05, P < 0.0001$). Comparing F1 scores from 10 classifier runs, post hoc Tukey tests reveal significant differences between the predictive powers of SPV versus CSP ($P < 0.001$), CSP versus Raw ($P < 0.001$), and SPV versus Raw features ($P < 0.001$). Maximum F1 scores were achieved with CSPs, while only close-to-chance classification performance is achieved with raw features.

Higher F1 scores were observed with the 1D-CNN for all features, revealing the superior learning ability of CNNs. A one-way ANOVA reveals the minimal impact of different features on 1D-CNN performance, even as CSP features achieve the highest F1 (0.86). The 2D- and 3D-CNNs achieve even higher F1 scores, conveying that the EEG image and movie representations are most effective for valence prediction. The 3D-CNN achieves the highest F1 score (0.91).

Figure 3A presents model sensitivity and specificity with PD data. Sensitivity denotes the true positive rate or proportion of HV samples classified correctly, while Specificity denotes the true negative rate or the proportion of correctly classified LV samples. For ML algorithms, a significantly higher mean specificity (0.94) is observed than sensitivity (0.67). A similar trend is observed for the 1D- and 2D-CNNs, with much higher specificity scores noted in both cases. Comparable mean specificity (0.95) and sensitivity (0.85) scores are noted, however, with the 3D-CNN. Overall, these trends convey reduced positive valence recognition with PD data.

Classification using HC data

F1 scores similar to or higher than PD are obtained on HC data for all features and methods (Table 4). The impact of features on ML performance is confirmed by an ANOVA test ($F(2, 27) = 1382.90, P < 0.0001$), with CSP features outperforming SPV and Raw features. Higher F1 scores are noted with the 1D-CNN, with all features performing similarly. The 2D- and 3D-CNNs perform better than the 1D-CNN, with the 3D-CNN achieving the best mean F1 score (0.93).

Table 4. Affect classification. F1 scores are of the form $\mu \pm \sigma$. The best results among all ML algorithms are reported in the ML column, corresponding to kNN and GNB (denoted using *).

		ML			1D-CNN				
Affect	Data	SPV	CSP	Raw	SPV	CSP	Raw	2D-CNN	3D-CNN
Valence	Full	0.78 ± 0.01	0.81 ± 0.01	0.55 ± 0.01	0.85 ± 0.01	0.82 ± 0.02	0.88 ± 0.11	0.89 ± 0.06	0.91 ± 0.05
	PD	0.75 ± 0.02	0.84 ± 0.01	0.55 ± 0.01	0.82 ± 0.09	0.86 ± 0.03	0.75 ± 0.17	0.86 ± 0.07	0.91 ± 0.07
	HC	0.81 ± 0.02	0.84 ± 0.01	0.56 ± 0.01	0.85 ± 0.08	0.88 ± 0.04	0.90 ± 0.11	0.91 ± 0.07	0.93 ± 0.05
Arousal	Full	0.92 ± 0.01	0.93 ± 0.00	0.76 ± 0.00*	0.95 ± 0.03	0.92 ± 0.01	0.95 ± 0.07	0.97 ± 0.02	0.97 ± 0.02
	PD	0.92 ± 0.01	0.93 ± 0.01	0.76 ± 0.00	0.96 ± 0.03	0.92 ± 0.02	0.94 ± 0.06	0.98 ± 0.02	0.98 ± 0.03
	HC	0.91 ± 0.01	0.94 ± 0.01	0.76 ± 0.00*	0.91 ± 0.05	0.95 ± 0.01	0.94 ± 0.07	0.94 ± 0.03	0.97 ± 0.02
Categorical emotions	Full	0.64 ± 0.01	0.72 ± 0.01	0.18 ± 0.01	0.76 ± 0.10	0.69 ± 0.02	0.77 ± 0.22	0.82 ± 0.09	0.83 ± 0.09
	PD	0.62 ± 0.02	0.76 ± 0.01	0.17 ± 0.01*	0.77 ± 0.08	0.80 ± 0.05	0.81 ± 0.21	0.84 ± 0.10	0.88 ± 0.09
	HC	0.68 ± 0.02	0.74 ± 0.02	0.20 ± 0.01	0.76 ± 0.09	0.76 ± 0.03	0.78 ± 0.22	0.86 ± 0.08	0.90 ± 0.07

PD versus HC F1 comparison

Figure 3B compares the valence F1 scores obtained with PD and HC data over all models, with CSP scores plotted for the ML and 1D-CNN methods. While identical scores were achieved on PD and HC data employing ML methods, marginally higher F1 scores were noted on the HC data for the 1D-, 2D-, and 3D-CNNs. Overall, better classification is achieved with the HC rather than the PD data. Figure 3C compares sensitivity for the PD and HC data across models. With ML algorithms, the sensitivity on PD data (0.67) is significantly lower than on HC data (0.73) as per a t test ($t(18) = 5.39, P < 0.0001$). Lower sensitivity scores are again noted on PD data with the 1D-, 2D-, and 3D-CNNs even if the differences are insignificant. Cumulatively, our result trends reveal lower valence recognition sensitivity to PD EEG data.

Arousal classification

To examine arousal perception in PD, we grouped the anger, disgust, fear, happiness, and surprise data in the high arousal (HA) category, and samples labeled sadness as constituting the low arousal (LA) category as in the circumplex model [10]. The HA:LA class ratio within PD, HC, and full data is, thus, 5:1.

Table 4 also presents arousal classification results with full, PD, and HC data. Evidently, similar F1 scores were achieved for these subsets. We again compare PD versus HC results. The arousal classification results of all ML algorithms on PD, HC, and full data are shown in Table S3.

Classification using PD data

The impact of the descriptors on ML classification performance is confirmed by a one-way ANOVA ($F(2, 27) = 1332.94$, $P < 0.0001$). Raw features performed worst (F1 = 0.76), while CSP features achieved optimal arousal prediction (F1 = 0.93). Significant F1 score differences were noted via a Tukey test for CSP versus SPV ($P < 0.005$), CSP versus Raw ($P < 0.001$), and SPV versus Raw ($P < 0.001$). Higher F1 scores were obtained for the 1D-CNN with spectral features performing best, even if the differences among descriptors were not significant. The 2D- and 3D-CNN models achieved an identical, near-ceiling F1 of 0.98.

Figure 4A presents specificity (LA classification rates) and sensitivity (HA classification rate) scores for the PD data. Significantly higher sensitivity (0.97) than specificity (0.76) was observed for the ML algorithms ($P < 0.0001$). This trend repeated for the 1D-CNN ($P < 0.0001$) and 3D-CNN (sensitivity = 0.98 > specificity = 0.91 with $P < 0.01$), while comparable measures were achieved for the 2D-CNN. Overall, higher sensitivity than specificity was achieved on PD data with the different models.

Classification using HC data

CSP features produced a maximum F1 score of 0.93 with ML methods on HC data. The 1D-CNN achieved a higher F1 of 0.95 with CSP features, but all features performed comparably. F1 scores of 0.94 and 0.97 were achieved with the 2D- and the 3D-CNN, respectively, revealing that the spectral EEG image and movie descriptors effectively encode emotion information.

PD versus HC F1 score comparison

The F1 scores achieved with PD and HC data are presented in Fig. 4B, with CSP results shown for the ML and 1D-CNN methods. Very similar F1 scores were found on PD and HC data for ML algorithms. The 1D-CNN achieved a much higher score with HC data ($P < 0.0001$), while the trend reversed for the 2D-CNN (PD F1 = 0.98 > HC F1 = 0.94, with $P < 0.05$). Similar F1 scores with PD and HC data were again noted for the 3D-CNN. Overall, result trends point to the lack of differences between EEG embeddings of the PD and HC cohorts with respect to arousal.

Categorical emotion classification

Since reduced HV-LV discriminability was noted in the PD patients, we explored categorical emotion recognition and the nature of misclassifications with PD versus HC data. A uniform distribution of 240 samples/emotion is available in this setting.

Multiclass emotion classification results across models with full, PD, and HC data are shown in Table 4. An equal number of samples were available for the sadness, happiness, fear, disgust, surprise, and anger emotion classes for both PD and HC subjects. F1 scores averaged over all emotion classes are shown. For most

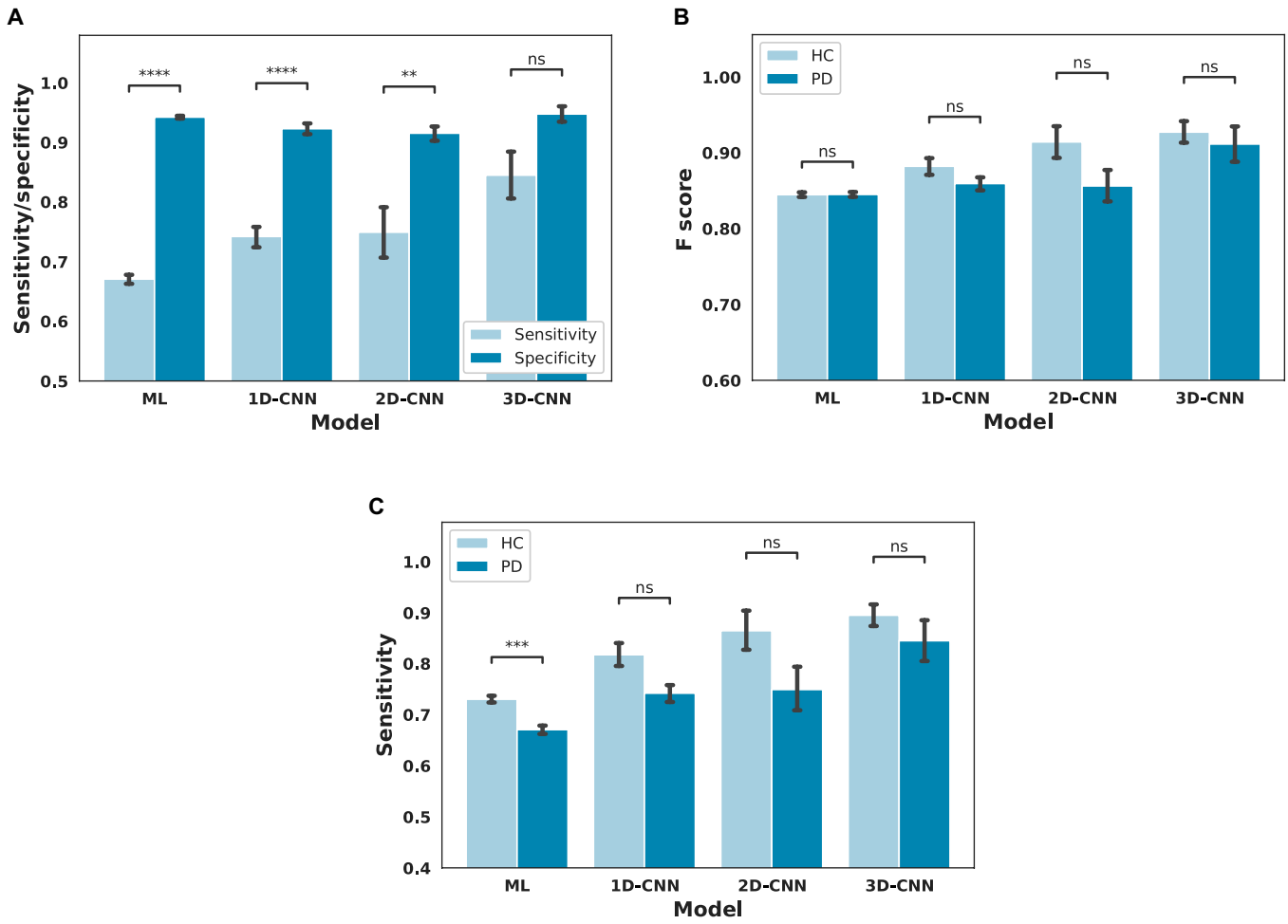

Fig. 3. Binary valence classification. (A) Sensitivity and specificity with PD data across various models. (B) F1 scores with PD and HC data across various models. (C) Sensitivity on PD and HC data across various models. Error bars denote standard error of mean (SEM). ****$P < 0.0001$, ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, and ns $P > 0.05$ as per a Tukey honestly significant difference (HSD) test.

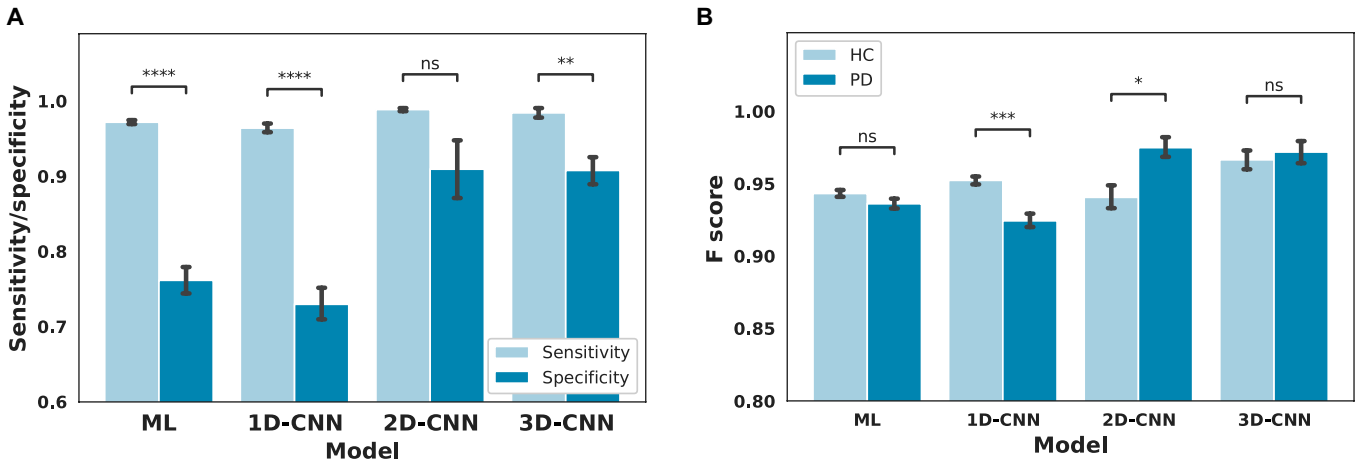

Fig. 4. Binary arousal classification. (A) Sensitivity and specificity with PD data across models. (B) F1 on PD and HC data across models. Error bars denote SEM. ****, ***, **, *, and ns respectively imply $P < 0.0001$, $P < 0.001$, $P < 0.01$, $P < 0.05$, and $P > 0.05$ as per a Tukey HSD test.

conditions, the highest scores were achieved with HC data, while the lowest scores were achieved with full data. The multiclass emotion classification results of all ML algorithms on PD, HC, and full data are shown in Table S4.

Classification with PD data

As with valence and arousal, EEG features significantly impacted ML performance as per a one-way ANOVA ($F(2, 27) = 4113.91$, $P < 0.0001$). CSP features produced the best F1 score (0.76),

significantly outperforming SPV and Raw features ($P < 0.001$) as per post hoc Tukey tests. For the 1D-CNN, raw EEG achieved the best F1 score (0.81), which was marginally superior to that of CSP and SPV features. Higher mean F1 scores of 0.84 and 0.88 were achieved with the 2D- and 3D-CNN, conveying accurate emotion recognition with PD data.

Figure 5 depicts emotion-specific F1 scores obtained on the PD and HC data across models, with CSP results presented for the ML and 1D-CNN models. For the PD data, the ML methods produced the highest and lowest F1 scores for sadness (F1 = 0.81) and surprise (F1 = 0.71), respectively, and a significant variation in F1 scores for different emotions was found per one-way ANOVA ($F(5, 54) = 8.68, P < 0.0001$). A significant effect of emotions on F1 scores was also noted for the 1D-, 2D-, and 3D-CNNs ($P < 0.05$ in all cases). Sadness was easiest to recognize with all 3 models (F1 = 0.87, 0.93, and 0.94 for the 1D-, 2D-, and 3D-CNN, respectively), while disgust (F1 = 0.78), surprise (F1 = 0.80), and fear (F1 = 0.85) were recognized worst by the 1D-, 2D-, and 3D-CNN, respectively.

Classification with HC data

Trends similar to those in the PD results were observed with the HC data. CSP features produced the highest score (F1 = 0.74) with the ML methods, significantly outperforming SPV and Raw features ($P < 0.001$ for both comparisons). All features performed comparably with the 1D-CNN, while the EEG image and movie descriptors produced mean F1 scores of 0.86 and 0.90, respectively, via the 2D- and 3D-CNN.

Misclassification analyses

Hitherto, (a) sensitivity–specificity analyses have shown lower recognition rates for HV emotions, and (b) emotion-specific results convey that surprise, disgust, and fear have often been confounded with other emotions. We further examined misclassifications with PD and HC data to discover any underlying patterns. Figure 6 depicts the maximum misclassification rate and most mispredicted label per model and emotion class. For instance, the first row shows that the sadness PD samples are often misclassified as happiness by the best-performing ML, 1D-, 2D-, and 3D-CNN models, with the misclassification rates specified in brackets. For HC data (2nd row), sadness is respectively mislabeled as happiness, anger, and fear. We note that:

- The happiness and surprise high-valence emotions are most commonly mislabeled as low-valence emotions, namely, sadness, fear, and anger for both PD and HC data. As per Fig. 6, misclassification rates are slightly higher with PD than HC data.
- Among low-valence emotions, sadness is consistently predicted as happiness with PD data. Conversely on HC data, sadness is often confounded with other low-valence emotions such as fear and anger.
- Fear and disgust are often misclassified with both PD and HC data. On HC data, fear is frequently confounded with

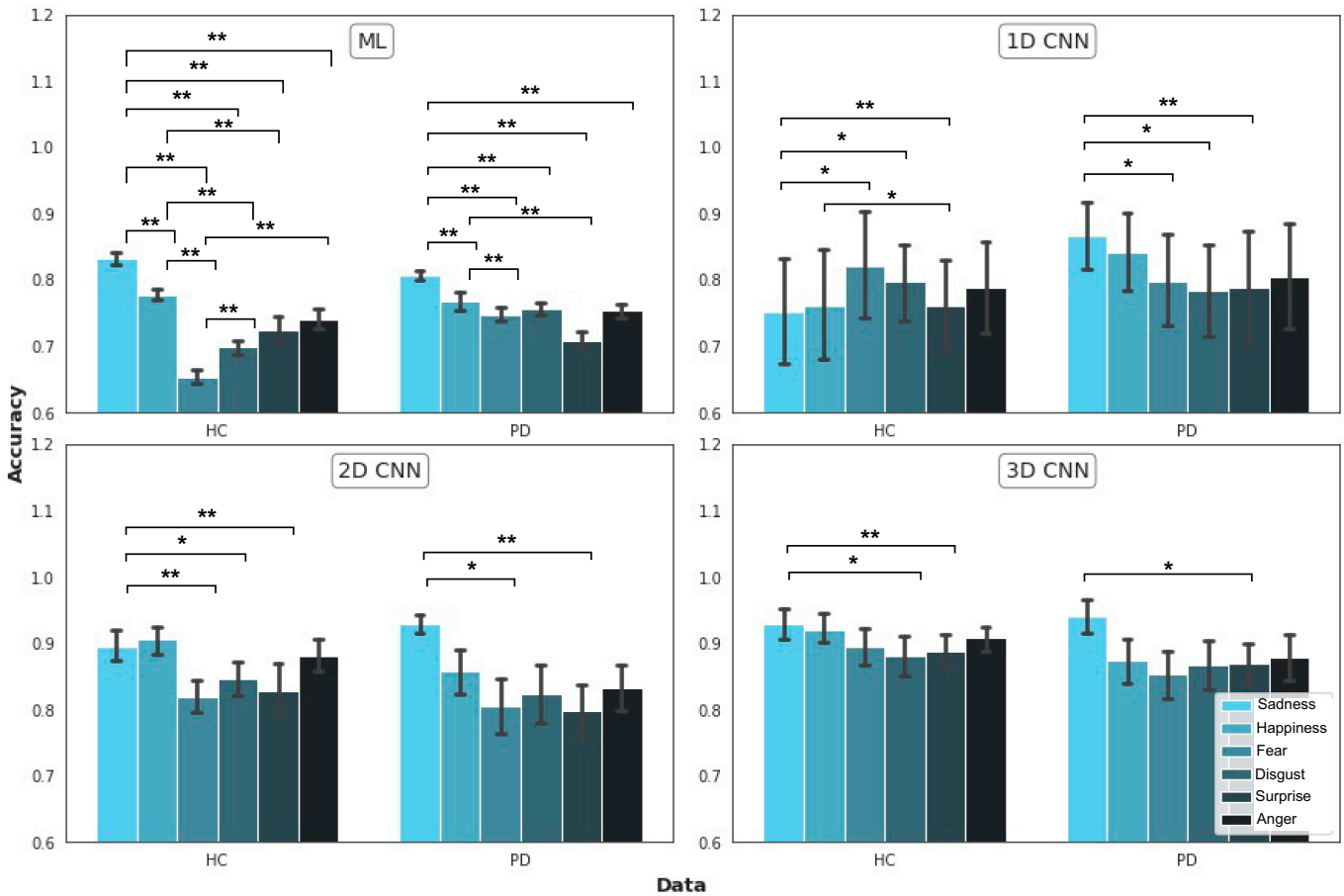

Fig. 5. Emotion-wise F1 scores on HC (left) and PD data (right) across models. Error bars denote the standard error of mean. ****, ***, **, *, and ns respectively imply $P < 0.0001$, $P < 0.001$, $P < 0.01$, $P < 0.05$, and $P > 0.05$ as per a Tukey HSD test.

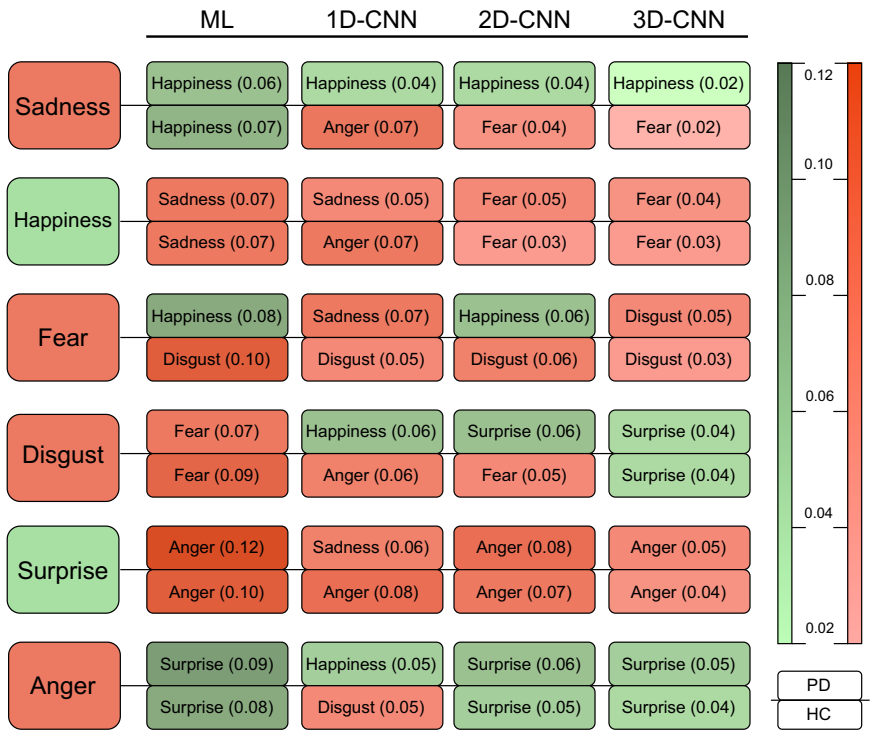

Fig. 6. Misclassifications on PD and HC emotion data across models. First column denotes the actual emotion, while other columns denote predicted emotion with the model specified on top. Upper and lower row pairs respectively denote emotions predicted with PD and HC data, and values in parentheses denote misclassification rates. Green and red colors are used to code positive and negative valence emotions.

disgust, and disgust with fear and anger. With PD data, however, disgust is often confused with happiness and surprise, and fear with happiness.

• From the above, one can infer a greater propensity to confound with opposite-valence emotions on PD data. Overall, misclassification trends results convey that valence-related differences are not effectively encoded in PD EEG responses.

PD versus HC classification

The above sections reveal some differences in the emotional EEG characteristics of the PD and HC groups. We examined whether the emotional EEG responses were discriminable for PD versus HC classification. To this end, we attempted PD recognition from both emotion-specific and emotion-agnostic (or full) EEG data acquired during the routine task of audiovisual media consumption. Given that a balanced class distribution is available in this setting (an equal number of PD and HC samples are available as part of the full as well as emotion-specific data), we further evaluated whether individual-specific rather than group-specific signatures impacted PD recognition performance by employing both the 10FCV and leave-one-out cross-validation (LOOCV) procedures on the best-performing 2D-CNN model (see Table 5).

Emotion-agnostic and specific PD recognition results are presented in Table 5. Empirical results show that accurate PD recognition is achieved even with limited and emotion-specific EEG data. Near-ceiling F1s are noted with the 2D-CNN, implying that PD and HC emotional responses are highly discriminable upon learning from only a few training samples. Notably, excellent discriminability between PD versus HC EEG encodings is confirmed by the excellent (even if marginally lower) F1 scores achieved with the LOOCV data-split strategy, where subject-specific data are held out and evaluated on the trained model. This observation indicates that the influence of subject-specific signatures is marginal for PD recognition from emotional EEG signals. The best recognition rates with the 10FCV (F1 of 0.99 with sadness and fear) and LOOCV (peak F1 of 0.98 with anger) are noted with negative valence emotions, mirroring prior findings concerning PD perception of negative emotions [25,27,28].

Further examining sensitivity and specificity measures, balanced recognition of PD and HC classes across models was achieved mainly for the negative disgust, fear, and sadness emotions. That negative emotions best reflect PD impairments has been observed in prior studies [27,28]. Focusing on features and models, spectral features achieved the best results with ML methods. With the 1D-CNN, superior F1 scores were achieved with all features even if no clear trends were discernible. Similar to emotion recognition, the 2D-CNN again achieved optimal PD recognition, demonstrating that the EEG movie features optimally encode PD-related emotional differences.

Discussion summary

Valence

We examined PD valence perception, since valence is a fundamental emotional attribute [57]. While prior studies have found valence-related differences between PD and HC groups via their explicit responses to visual [27], verbal [30], and textual [26] stimuli, we differently examined implicit emotional EEG responses to this end. As per the results in Table 4 and Fig. 3, valence classification trends with multiple features and methods convey reduced performance and lower sensitivity on PD data in general. Therefore, PD data exhibit lower valence discriminability and sensitivity.

Table 5. PD versus HC results. F1 scores (equals accuracy for balanced dataset) are of the form $\mu \pm \sigma$. Best ML results are achieved with the kNN and GNB classifiers. GNB results are denoted with a * symbol. Best F1 scores are shown in bold.

	ML			1D-CNN			2D-CNN	
Data	SPV	CSP	Raw	SPV	CSP	Raw	10FCV	LOOCV
Full	0.97 ± 0.01	0.88 ± 0.01	0.66 ± 0.01	$\mathbf{0.99 \pm 0.01}$	0.87 ± 0.02	0.91 ± 0.07	$\mathbf{0.99 \pm 0.01}$	0.96 ± 0.09
Sadness	0.97 ± 0.01	0.92 ± 0.01	0.58 ± 0.03	0.97 ± 0.02	0.94 ± 0.02	0.96 ± 0.06	$\mathbf{0.99 \pm 0.01}$	0.95 ± 0.12
Happiness	0.97 ± 0.01	0.90 ± 0.02	0.59 ± 0.02*	0.91 ± 0.08	0.92 ± 0.02	0.95 ± 0.09	0.98 ± 0.02	0.97 ± 0.06
Fear	0.96 ± 0.01	0.93 ± 0.02	0.61 ± 0.03*	0.96 ± 0.02	0.95 ± 0.02	0.89 ± 0.15	$\mathbf{0.99 \pm 0.01}$	0.95 ± 0.07
Disgust	0.96 ± 0.02	0.91 ± 0.02	0.59 ± 0.01*	0.98 ± 0.03	0.93 ± 0.03	0.93 ± 0.09	0.98 ± 0.02	0.95 ± 0.14
Surprise	0.95 ± 0.01	0.91 ± 0.03	0.60 ± 0.02*	0.92 ± 0.03	0.94 ± 0.02	0.94 ± 0.10	0.97 ± 0.03	0.94 ± 0.16
Anger	0.96 ± 0.01	0.92 ± 0.01	0.59 ± 0.03*	0.98 ± 0.01	0.94 ± 0.03	0.87 ± 0.14	0.98 ± 0.02	0.98 ± 0.03

Lower PD sensitivity is consistent with the findings in [29], where PD patients are found to have deficits in processing both positive and negative emotions. Dysfunction of the basal ganglia thalamocortical circuits in PD patients impairs their general emotional valence recognition [2]. With respect to models and features, CNNs expectedly achieved higher F1 scores than ML methods (see Table 4), confirming that they can efficiently learn spatiotemporal EEG patterns. CSP features predominantly achieve the best scores, and their utility in EEG-based analysis is well known [48]. The EEG movie descriptor optimally encodes spatiotemporal patterns in spectral EEG.

Arousal

With respect to arousal, Miller et al. [13] observed muted reactivity or fewer startled eye blinks from PD patients to high-arousal, low-valence aversive pictures. Similar findings were reported in [11,26], where PD patients showed deficits in recognizing emotions from lexical, prosody, and facial cues. While these findings are based on implicit EMG data and/or explicit self-ratings, our inferences are based on EEG classification patterns. Given the identical distributions for the PD and HC classes and corresponding F1 comparisons, our results convey that very comparable F1 scores are obtained for the PD and HC groups across models and features, conveying little difference between their EEG encodings with respect to arousal.

Performance of models and features

Sensitivity and specificity scores observed for PD data also convey an interesting trend. Given the imbalanced class proportions for both the valence and arousal conditions, significant sensitivity versus specificity disparities are noted particularly for the ML and 1D-CNN models. However, these differences become less conspicuous for the 2D- and 3D-CNN models, conveying that they are able to efficiently learn minority-class representations. Regarding algorithms and features, F1 scores gradually improve while advancing from the ML algorithms to the 3D-CNN. CSP features perform best with ML algorithms, but mixed results are observed for the 1D-CNN. The EEG image and movie descriptors achieve a maximum and identical F1 score of 0.98 for arousal, with the 3D-CNN producing the maximum F1 of 0.93 for valence, showing their efficacy in learning emotional EEG representations.

Categorical emotion recognition

We creatively examined categorical emotional classification using EEG data from both PD and HC participants. Prior studies on PD emotional perception typically examined facial expression recognition tasks [12,26,28] or studied physiological signals along with self-assessment reports to understand emotional deficits [32,34]. These studies observe PD impairment in recognizing negative emotions such as sadness, fear, anger, and disgust.

We performed categorical emotion recognition to better understand which emotions are recognized better/worse with PD and HC data. Our results in the "Classification with PD data" section reveal that while sadness and happiness were well recognized with both PD and HC data, fear, disgust, and surprise were poorly recognized with the PD EEG data. Disturbances in the orbitofrontal cortex and the anterior cingulate cortex, which are active in negative emotion processing, can be attributed to these deficits [29]. We then studied the nature of misclassifications for each emotion class. Misclassification results in Fig. 3 show frequent confounds among opposite-valence emotions with PD data, indicating weaker valence encodings in emotional PD EEG responses.

With respect to features and models, the trends are consistent with valence and arousal classification. We generally observe a steady increase in emotion-specific and overall F1 scores as we progress from the classical ML methods to the 3D-CNN. The CSP and raw EEG features produce the best performance with the ML and 1D-CNN approaches, respectively. The 2D- and 3D-CNN models, however, tend to achieve higher F1 scores, implying that spatiotemporal spectral EEG patterns best encode emotional information.

PD versus HC classification

To our knowledge, only 1 study [58] has performed PD recognition from emotional EEG signals. In it, a mean accuracy of 87.9% was achieved employing ML classifiers and a 10FCV data-split strategy (see Table 1). Others [6,22] have performed PD versus HC classification from resting-state EEG signals, which is an ecologically invalid setting requiring a highly controlled environment for EEG acquisition. In this regard, we attempted PD recognition from both emotion-specific and emotion-agnostic (or full) EEG data acquired during the routine task of audiovisual media consumption. Recent studies [59–61] observe that individual-specific, rather than emotion-specific, signatures are utilized by ML models to achieve superior recognition on small

datasets, and that recognition performance is adversely impacted when subject-independent train and test sets are used for evaluation. To examine if this observation holds for PD recognition, we employed both the 10FCV and LOOCV data-split strategies to train and evaluate the best-performing 2D-CNN model.

The empirical results presented in Table 5 show that accurate PD recognition is achieved even with emotion-specific EEG data. Near-perfect F1s are noted with the 2D-CNN, implying that PD and HC emotional responses are highly discriminable upon learning from only a few training samples. Furthermore, PD versus HC discriminability is not influenced by individual signatures as conveyed by the excellent recognition performance achieved with the LOOCV data-split strategy, where the test subject data are held out from training. While the best recognition rates with the 10FCV procedure are obtained for the sadness and fear emotions, a peak F1 score of 0.98 is achieved for anger with LOOCV. The fact that maximum PD versus HC discriminability can be noted for negative emotions mirrors similar findings in [12,26–28].

Further examining sensitivity and specificity measures, balanced recognition of PD and HC classes across models was achieved mainly for the negative disgust, fear, and sadness emotions, reinforcing that negative emotions best reflect PD impairments [27,28]. Focusing on features and models, spectral features achieved the best results with ML methods. With the 1D-CNN, superior F1 scores were achieved with all features even if no clear trends were discernible. Similar to emotion recognition, 2D-CNN again achieved optimal PD recognition, demonstrating that the EEG movie features optimally encode PD-related emotional differences.

Conclusion

PD patients may often have difficulty expressing their emotions and internal feelings in real life owing to (a) PD effects especially in the advanced stages of the disease and (b) the effect of associated medications. Given these limitations, an assistive and sustainable diagnostic tool based on noninvasive detection of emotional disturbances can facilitate treatment and help improve life quality for PD patients. While many studies identify PD-related impairments based on the patients' explicit and implicit responses [11,13,62], or cognitive dissimilarities based on resting-state EEG [4,6,22], we differently examined emotional EEG responses to perform both emotion and PD recognition.

While studies examining facial behavior and resting-state EEG [4] typically derive their findings based on statistical patterns observed for the PD and HC groups, our inferences are entirely derived from classification patterns. Interesting trends and similarities with prior work are revealed in our analyses. Dimensional emotion recognition experiments conveyed reduced discriminability with PD EEG data, while arousal-related differences vis-à-vis the HC group were not apparent. Furthermore, categorical emotion recognition results revealed that disgust, fear, and surprise were associated with low recognition rates on PD data, while sadness was well recognized. Mislabeling analyses showed frequent confounds among opposite-valence emotions with PD data, but not with HC data. Reduced recognition of low-valence emotions and confounds noted with positive emotions mirror with deficits noted in the perception of these emotions from pictorial [13] and prosodic stimuli [11].

Given some differences in emotion perception between the PD and HC groups, we then examined if the PD versus HC emotional responses were discriminable, and if this discriminability differed across emotions. The empirical results revealed that differences were apparent for both emotion-specific and emotion-agnostic data, with high F1 scores achieved in all conditions. Here, again, the maximum recognition performance was achieved for negative emotions; also, very similar sensitivity and specificity rates across models were noted for negative emotions such as disgust, sadness, and fear, reinforcing that the PD versus HC encodings were most discriminable for these emotions.

With respect to features and models, CSPs considerably outperformed spectral features with ML models for emotion recognition. Conversely, spectral descriptors outperformed CSPs for PD versus HC classification. The efficacy of spectral features for isolating PD characteristics has been observed in prior studies [63]. No single feature performed best with the 1D-CNN, even if the 1D-CNN consistently outperformed classical ML methods. The 2D- and 3D-CNN models consistently achieved optimal recognition performance, conveying that spectral spatiotemporal models best encode EEG patterns as noted in [20,56].

Our key finding is that both emotion and PD recognition can be reliably performed from EEG responses passively compiled during audiovisual stimulus viewing. Given that we effortlessly interact with media routinely, EEG signals can be captured easily over longer time intervals as compared to resting-state EEG, which can practically be acquired only over short episodes. Also, while many EEG differences between PD and HC groups have been noted from resting-state analysis, the exact relation between EEG and motor symptoms is unknown [4].

Most empirical results presented in this paper are generated via a 10FCV strategy, given the small dataset size and imbalanced class distribution. Nevertheless, the LOOCV procedure adopted for the PD versus HC classification problem reveals that the EEG encodings are relatively robust to individual idiosyncrasies. A limitation of our study is that we analyze data compiled from a limited number of PD subjects with only mild-to-moderate disease severity (Hoehn and Yahr scale [64] of 1 to 3). Future work will also focus on severity levels 4 and 5. While perceptual differences between PD versus HC subjects were captured via classification results, an assistive diagnostic tool should also be able to provide explanations behind decision-making. Future work will focus on generating explanatory predictions, building on recent work [65] in this regard.

Acknowledgments

Funding: This research was supported partially by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP190101294).
Author contributions: R.S. and M.M. conceived the idea and designed the experiments, with the help of I.R. and R.G. R.P. and S.N. wrote the code and performed the experiments. All authors contributed to the analysis of the results and the writing of the manuscript.
Competing interests: The authors declare that they have no competing interests.

Data Availability

The raw EEG data of PD patients and HCs used in this study are sensitive medical data collected from Hospital Kebangsaan Malaysia, Kuala Lumpur. Reasonable requests for accessing the EEG data used in this study can be directed to M.M.

(email: m.murugappan@kcst.edu.kw). The source code to generate features is available on GitHub at https://github.com/ravikiranrao/EEG-Parkinson.

Supplementary Materials

Dataset
Results
Tables S1 to S5

References

1. Dauer W, Przedborski S. Parkinson's disease: Mechanisms and models. *Neuron*. 2003;39(6):889–909.
2. Trojano L, Papagno C. Cognitive and behavioral disorders in Parkinson's disease: An update. II: Behavioral disorders. *Neurol Sci*. 2018;39(1):53–61.
3. Chopade P, Chopade N, Zhao Z, Mitragotri S, Liao R, Chandran SV. Alzheimer's and Parkinson's disease therapies in the clinic. *Bioeng Transl Med*. 2023;8(1): Article e10367.
4. Wang Q, Meng L, Pang J, Zhu X, Ming D. Characterization of EEG data revealing relationships with cognitive and motor symptoms in Parkinson's disease: A systematic review. *Front Aging Neurosci*. 2020;12:373.
5. Chaturvedi M, Hatz F, Gschwandtner U, Bogaarts JG, Meyer A, Fuhr P, Roth V. Quantitative EEG (QEEG) measures differentiate Parkinson's disease (PD) patients from healthy controls (HC). *Front Aging Neurosci*. 2017;9.
6. Oh SL, Hagiwara Y, Raghavendra U, Yuvaraj R, Arunkumar N, Murugappan M, Acharya UR. A deep learning approach for Parkinson's disease diagnosis from EEG signals. *Neural Comput Appl*. 2020;32:10927–10933.
7. Koelstra S, Muhl C, Soleymani M, Jong-Seok Lee, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I. DEAP: A database for emotion analysis using physiological signals. *IEEE Trans Affect Comput*. 2012;3(1):18–31.
8. Yuvaraj R, Murugappan M, Acharya UR, Adeli H, Ibrahim NM, Mesquita E. Brain functional connectivity patterns for emotional state classification in Parkinson's disease patients without dementia. *Behav Brain Res*. 2016;298:248–260.
9. Abadi MK, Subramanian R, Kia SM, Avesani P, Patras I, Sebe N. DECAF: MEG-based multimodal database for decoding affective physiological responses. *IEEE Trans Affect Comput*. 2015;6(3):209–222.
10. Russell JA. A circumplex model of affect. *J Pers Soc Psychol*. 1980;39(6):1161–1178.
11. Paulmann S, Pell MD. Dynamic emotion processing in Parkinson's disease as a function of channel availability. *J Clin Exp Neuropsychol*. 2010;32(8):822–835.
12. Clark US, Neargarder S, Cronin-Golomb A. Specific impairments in the recognition of emotional facial expressions in Parkinson's disease. *Neuropsychologia*. 2008;46(9):2300–2309.
13. Miller K, Okun M, Marsiske M, Fennell E, Bowers D. Startle reflex hyporeactivity in Parkinson's disease: An emotion-specific or arousal-modulated deficit? *Neuropsychologia*. 2009;47(8–9):1917–1927.
14. RJR B. Facial expressions, their communicatory functions and neuro–cognitive substrates. *Philos Trans R Soc Lond B Biol Sci*. 2003;358(1431):561–572.
15. Narayana S, Subramanian R, Radwan I, Goecke R. To improve is to change: Towards improving mood prediction by learning changes in emotion. Paper presented at: Companion Publication of the 2022 International Conference on Multimodal Interaction; 2022; Bengaluru, India.
16. Subramanian R, Wache J, Abadi MK, Vieriu RL, Winkler S, Sebe N. ASCERTAIN: Emotion and personality recognition using commercial sensors. *IEEE Trans Affect Comput*. 2018;9(2):147–160.
17. Shukla A, Gullapuram SS, Katti H, Kankanhalli M, Winkler S, Subramanian R. Recognition of advertisement emotions with application to computational advertising. *IEEE Trans Affect Comput*. 2022;13(2):781–792.
18. Petrantonakis PC, Hadjileontiadis LJ. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Trans Affect Comput*. 2010;1(2):81–97.
19. Balconi M, Lucchiari C. Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis. Int'l. *J Psychophysiol*. 2008;67(1): 41–46.
20. Bashivan P, Rish I, Yeasin M, Codella N. Learning representations from EEG with deep recurrent-convolutional neural networks. Paper presented at: International Conference on Learning Representations; 2016; San Juan, Puerto Rico.
21. Ai Q, Liu Q, Meng W, Xie SQ. Chapter 6—EEG-based brain intention recognition. In: Ai Q, Liu Q, Meng W, Xie SQ, editors. *Advanced rehabilitative technology*. Cambridge (MA): Academic Press; 2018. p. 135–166.
22. Bhurane AA, Dhok S, Sharma M, Yuvaraj R, Murugappan M, Acharya UR. Diagnosis of Parkinson's disease from electroencephalography signals using linear and self-similarity features. *Expert Syst*. 2019;39(7):Article e12472.
23. Mantri S, Wood S, Duda JE, Morley JF. Comparing self-reported and objective monitoring of physical activity in Parkinson disease. *Parkinsonism Relat Disord*. 2019;67:56–59.
24. Hsu YT, Lai HY, Chang YC, Chiou SM, Lu MK, Lin YC, Liu YL, Chen CC, Huang HC, Chien TF, et al. The role of the sub-thalamic nucleus in the preparation of volitional movement termination in Parkinson's disease. *Exp Neurol*. 2012;233(1):253–263.
25. Gray HM, Tickle-Degnen L. A meta-analysis of performance on emotion recognition tasks in Parkinson's disease. *Neuropsychology*. 2010;24(2):176.
26. Kan Y, Kawamura M, Hasegawa Y, Mochizuki S, Nakamura K. Recognition of emotion from facial, prosodic and written verbal stimuli in Parkinson's disease. *Cortex*. 2002;38(4): 623–630.
27. Baggio HC, Segura B, Ibarretxe-Bilbao N, Valldeoriola F, Marti MJ, Compta Y, Tolosa E, Junqué C. Structural correlates of facial emotion recognition deficits in Parkinson's disease patients. *Neuropsychologia*. 2012;50(8):2121–2128.
28. Narme P, Bonnet AM, Dubois B, Chaby L. Understanding facial emotion perception in Parkinson's disease: The role of configural processing. *Neuropsychologia*. 2011;49(12):3295–3302.
29. Lin CY, Tien YM, Huang JT, Tsai CH, Hsu LC. Degraded impairment of emotion recognition in Parkinson's disease extends from negative to positive emotions. *Behav Neurol*. 2016;2016:9287092.
30. Jin Y, Mao Z, Ling Z, Xu X, Zhang Z, Yu X. Altered emotional recognition and expression in patients with Parkinson's disease. *Neuropsychiatr Dis Treat*. 2017;13:2891–2902.

31. Bilalpur M, Kia SM, Chawla M, Chua TS, Subramanian R. Gender and emotion recognition with implicit user signals. Paper presented at: International Conference on Multimodal Interaction; 2017; Glasgow, UK.
32. Wabnegger A, Ille R, Schwingenschuh P, Katschnig-Winter P, Kögl-Wallner M, Wenzel K, Schienle A. Facial emotion recognition in Parkinson's disease: An fMRI investigation. *PLOS ONE*. 2015;10(8):Article e0136110.
33. Moonen AJ, Weiss PH, Wiesing M, Weidner R, Fink GR, Reijnders JSAM, Weber WM, Leentjens AFG. An fMRI study into emotional processing in Parkinson's disease: Does increased medial prefrontal activation compensate for striatal dysfunction? *PLOS ONE*. 2017;12(5): Article e0177085.
34. Wu P, Gonzalez I, Patsis G, Jiang D, Sahli H, Kerckhofs E, Vandekerckhove M. Objectifying facial expressivity assessment of Parkinson's patients: Preliminary study. *Comput Math Methods Med*. 2014;2014:427826.
35. Tessitore A, Hariri AR, Fera F, Smith WG, Chase TN, Hyde TM, Weinberger DR, Mattay VS. Dopamine modulates the response of the human amygdala: A study in Parkinson's disease. *J Neurosci*. 2002;22(20):9099–9103.
36. Mirowski P, Madhavan D, LeCun Y, Kuzniecky R. Classification of patterns of EEG synchronization for seizure prediction. *Clin Neurophysiol*. 2009;120(11):1927–1940.
37. Qureshi MNI, Oh J, Lee B. 3D-CNN based discrimination of schizophrenia using resting-state fMRI. *Artif Intell Med*. 2019;98:10–17.
38. Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, Johnson HJ, Paulsen JS, Turner JA, Calhoun VD. Deep learning for neuroimaging: A validation study. *Front Neurosci*. 2014;8:229.
39. Yuvaraj R, Murugappan M, Ibrahim NM, Omar MI, Sundaraj K, Mohamad K, Palaniappan R, Satiyan M. Emotion classification in Parkinson's disease by higher-order spectra and power spectrum features using EEG signals: A comparative study. *J Integr Neurosci*. 2014;13(1):89–120.
40. Yuvaraj R, Murugappan M, Ibrahim NM, Sundaraj K, Omar MI, Mohamad K, Palaniappan R. Optimal set of EEG features for emotional state classification and trajectory visualization in Parkinson's disease. *Int J Psychophysiol*. 2014;94(3):482–495.
41. Yuvaraj R, Murugappan M. Hemispheric asymmetry non-linear analysis of EEG during emotional responses from idiopathic Parkinson's disease patients. *Cogn Neurodyn*. 2016;10(3): 225–234.
42. Lee S, Hussein R, McKeown MJ. A deep convolutional-recurrent neural network architecture for Parkinson's disease EEG classification. Paper presented at: 2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP); 2019; Ottawa, Canada.
43. Bradley MM, Lang PJ. *The International Affective Digitized Sounds: Affective ratings of sounds and instruction manual (Technical Report No. B-3)*. Gainesville (FL): University of Florida; 2007.
44. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): Technical manual and affective ratings. *NIMH CSEA*. 1997;1:3.
45. Jenkinson N, Brown P. New insights into the relationship between dopamine, beta oscillations and motor function. *Trends Neurosci*. 2011;34(12):611–618.
46. Kim J. Bimodal emotion recognition using speech and physiological changes. In: *Robust speech recognition and understanding*. Rijeka (Croatia): InTech; 2007. Vol. 265; p. 280.
47. Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J. Deep learning based electroencephalography analysis: A systematic review. *J Neural Eng*. 2019;16(5):Article 051001.
48. Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans Rehabil Eng*. 2000;8(4):441–446.
49. Müller-Gerking J, Pfurtscheller G, Flyvbjerg H. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clin Neurophysiol*. 1999;110(5):787–798.
50. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13:21–27.
51. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273–297.
52. John GH, Langley P. Estimating continuous distributions in Bayesian classifiers. arXiv. 2013. https://doi.org/10.48550/arXiv.1302.4964.
53. Breiman L, Friedman JH, Olshen RA, Stone C. *Classification and regression trees*. Monterey (CA): Wadsworth and Brooks/Cole; 1984.
54. Cohen J, Cohen P, West SG, Aiken LS. *Applied multiple regression/correlation analysis for the behavioral sciences.* New York (NY): Routledge; 2013.
55. Cramer JS. The early origins of the logit model. *Stud Hist Philos Sci C: Stud Hist Philos Biol Biomed Sci*. 2004;35(4):613–626.
56. Chugh K, Gupta P, Dhall A, Subramanian R. Not made for each other: Audio-visual dissonance-based deepfake detection and localization. Paper presented at: ACM International Conference on Multimedia; 2020; Seattle, WA, USA.
57. Shuman V, Sander D, Scherer KR. Levels of valence. *Front Psychol*. 2013;4:261.
58. Yuvaraj R, Murugappan M, Mohamed Ibrahim N, Sundaraj K, Omar MI, Mohamad K, Palaniappan R. Detection of emotions in Parkinson's disease using higher order spectral features from brain's electrical activity. *Biomed Signal Process Control*. 2014;14:108–116.
59. Sharma G, Pandey P, Subramanian R, Miyapuram KP, Dhall A. Neural encoding of songs is modulated by their enjoyment. Paper presented at: International Conference on Multimodal Interaction; 2022; Bengaluru, India.
60. Parameshwara R, Radwan I, Subramanian R, Goecke R. Examining subject-dependent and subject-independent human affect inference from limited video data. Paper presented at: International Conference on Automatic Face and Gesture Recognition; 2023; Puako, HI, USA.
61. Parameshwara R, Radwan I, Asthana A, Abbasnejad I, Subramanian R, Goecke R. Efficient labelling of affective video datasets via few-shot & multi-task contrastive learning. In: *Proceedings of the 31st ACM International Conference on Multimedia. MM '23*. New York (NY): Association for Computing Machinery; 2023. p. 6161–6170.
62. Smith MC, Smith MK, Ellgring H. Spontaneous and posed facial expression in Parkinson's disease. *J Int Neuropsychol Soc*. 1996;2(5):383–391.
63. Neufeld M, Blumen S, Aitkin I, Parmet Y, Korczyn A. EEG frequency analysis in demented and nondemented Parkinsonian patients. *Dementia*. 1994;5(1):23–28.
64. Hoehn M, Yahr M. Parkinsonism: Onset, progression, and mortality. *Neurology*. 1998;50(2):318–334.

65. Assaf R, Schumann A. Explainable deep neural networks for multivariate time series predictions. Paper presented at: International Joint Conference on Artificial Intelligence; 2019; Macao, China.
66. Jacobs DH, Shuren J, Bowers D, Heilman KM. Emotional facial imagery, perception, and expression in Parkinson's disease. *Neurology*. 1995;45(9):1696–1702.
67. Simons G, Ellgring H, Smith PM. Disturbance of spontaneous and posed facial expressions in Parkinson's disease. *Cognit Emot*. 2003;17:759–778.
68. Dar MN, Akram MU, YUVARAJ R, Khawaja SG, Murugappan M. EEG-based emotion charting for Parkinson's disease patients using convolutional recurrent neural networks and cross dataset learning. *Comput Biol Med*. 2022;144:Article 105327.