HIR
Healthcare Informatics Research

Table S1. Time, frequency, and cepstral domain extracted features using short-term and mid-term processing techniques [15,16]

Feature ID	Domain	Feature name	Description
1	Time domain	Zero Crossing Rate	The rate in which the sign of the signal changes during the frame.
2		Energy	The sum of squares of the signal values normalized by its frame length.
3		Entropy of Energy	The measure of abrupt changes in the energy level of the signal.
4	Frequency domain	Spectral Centroid	The center of gravity of the spectrum.
5		Spectral Spread	The second central moment of the spectrum.
6		Spectral Entropy	Entropy of the normalized spectral energies for a set of subframes.
7		Spectral Flux	The squared difference between the normalized magnitudes of the spectra of the two successive frames.
8		Spectral Rolloff	The frequency below which 90% of the magnitude distribution of the spectrum is concentrated.
9–21	Cepstral domain	MFCCs	Mel frequency cepstral coefficients (MFCC) form a cepstral representation where the frequency bands are not linear but distributed according to the mel-scale.
22–33	Frequency domain	Chroma Vector	A 12-element representation of the spectral energy where the bins represent the 12 equal-tempered pitch classes of western-type music.
34		Chroma Deviation	The std of the 12 chroma coefficients.

Table S2. Final dataset features

Feature ID	Feature name	Description
1	age	The age of the participant
2	gender	The gender of the participant: 1 for males 0 for females
3–36	features_mean	The mean of each of the 34 features, example: energy_mean spectral_centroid_mean
37–70	delta_features_mean	The delta of the mean of each of the 34 features, example: delta_energy_mean delta_spectral_centroid_mean
71–104	features_std	The std of each of the 34 features, example: energy_std spectral_centroid_std
105–138	delta_features_std	The delta of the std of each of the 34 features, example: delta_energy_std delta_pectral_centroid_std
139	class	The supervised learning class: 1 for PD 0 for control

PD: Parkinson disease.

Table S3. Selected subset of features using ANOVA and LASSO with various parameter values

Feature's rank	ANOVA	LASSO		
	Best 30 features	Best 11 features (C = 0.01)	Best 21 features (C = 0.02)	Best 33 features (C = 0.03)
1	age	gender	gender	gender
2	gender	age	age	age
3	energy_entropy_mean	energy_entropy_mean	energy_entropy_mean	energy_entropy_mean
4	mfcc_9_mean	spectral_spread_mean	spectral_spread_mean	spectral_spread_mean
5	spectral_spread_mean	mfcc_6_mean	mfcc_1_mean	mfcc_1_mean
6	mfcc_8_mean	mfcc_8_mean	mfcc_4_mean	mfcc_2_mean
7	mfcc_10_mean	mfcc_9_mean	mfcc_6_mean	mfcc_3_mean
8	energy_mean	mfcc_10_mean	mfcc_8_mean	mfcc_4_mean
9	mfcc_6_mean	mfcc_12_mean	mfcc_9_mean	mfcc_6_mean
10	spectral_centroid_mean	chroma_8_mean	mfcc_10_mean	mfcc_8_mean
11	energy_entropy_std	delta_spectral_spread_std	mfcc_11_mean	mfcc_9_mean
12	chroma_std_mean	-	mfcc_12_mean	mfcc_10_mean
13	chroma_8_mean	-	chrom_6_mean	mfcc_11_mean
14	delta_mfcc_7_std	-	chroma_8_mean	mfcc_12_mean
15	delta_energy_entropy_std	-	chroma_11_mean	chroma_5_mean
16	delta_mfcc_4_std	-	energy_entropy_std	chroma_6_mean
17	energy_std	-	mfcc_3_mean	chroma_8_mean
18	delta_mfcc_5_std	-	delta_spectral_spread_std	chroma_10_mean
19	delta_energy_std	-	delta_mfcc_3_std	chroma_11_mean
20	delta_spectral_spread_std	-	delta_mfcc_12_std	energy_entropy_std
21	delta_mfcc_6_std	-	delta_chroma_7_std	mfcc_3_std
22	zcr_mean	-	-	mfcc_8_std
23	mfcc_12_mean	-	-	mfcc_13_std
24	spectral_entropy_mean	-	-	chroma_8_std
25	mfcc_13_mean	-	-	delta_spectral_spread_std
26	chroma_8_std	-	-	delta_mfcc_2_std
27	mfcc_2_mean	-	-	delta_mfcc_3_std
28	mfcc_8_std	-	-	delta_mfcc_11_std
29	mfcc_6_std	-	-	delta_mfcc_12_std
30	mfcc_5_std	-	-	delta_chroma_3_std
31	-	-	-	delta_chroma_7_std
32	-	-	-	delta_chroma_9_std
33	-	-	-	delta_chroma_12_std

ANOVA: analysis of variance, LASSO: least absolute shrinkage and selection operator.