IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 2, 2021, accepted October 3, 2021, date of publication October 8, 2021, date of current version October 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3119035

Exploiting Spectral and Cepstral Handwriting Features on Diagnosing Parkinson's Disease

JUAN A. NOLAZCO-FLORES[1], (Senior Member, IEEE), MARCOS FAUNDEZ-ZANUY[2], V. M. DE LA CUEVA[1], (Member, IEEE), AND JIRI MEKYSKA[3]

[1]School of Engineering and Science, Tecnológico de Monterrey, Monterrey, Nuevo León 64849, Mexico
[2]Escola Superior Politecnica, TecnoCampus Mataro-Maresme, 08302 Mataro, Spain
[3]Department of Telecommunications and SIX Research Center, Brno University of Technology, 61600 Brno, Czechia

Corresponding authors: Juan A. Nolazco-Flores (jnolazco@tec.mx) and V. M. de la Cueva (vcueva@tec.mx)

This work was supported in part by MINECO Spanish Grant PID2020-113242RB-I00.

ABSTRACT Parkinson's disease (PD) is the second most frequent neurodegenerative disease associated with several motor symptoms, including alterations in handwriting, also known as PD dysgraphia. Several computerized decision support systems for PD dysgraphia have been proposed, however, the associated challenges require new approaches for more accurate diagnosis. Therefore, this work adds spectral and cepstral handwriting features to the already-used temporal, kinematic and statistics handwriting features. First, we calculate temporal and kinematic features using displacement; statistic features (*SF*) using displacement, and horizontal and vertical displacement; spectral (*SDF*) and cepstral (*CDF*) using displacement, horizontal and vertical displacement and pressure. Since the employed dataset (PaHaW) contains only 37 PD patients and 38 healthy control subjects (HC), then as the second step, we augment the percentage of the smaller training set to equal the larger. Next, we augment both classes to increase the training patient's data and added random Gaussian noise in all augmentations. Third, the most relevant features were selected using the modified fast correlation-based filtering method (mFCBF). Finally, autoML is employed to train and test more than ten plain and ensembled classifiers. Experimental results show that adding spectral and cepstral features to temporal, kinematics and statistics features highly improved classification accuracy to 98.57%. Our proposed model, with lower computational complexities, outperforms conventional state-of-the-art models for all tasks, which is 97.62%.

INDEX TERMS Parkinson's disease, dysgraphia, online handwriting, feature extraction, data augmentation, autoML.

I. INTRODUCTION

Biometrics can be used for e-security and e-health [1] and can be grouped based on two traits. Morphological biometrics, such as fingerprints or eye pupils, use direct measurements of the physical traits of the human body [2], [3]. Behavioral biometrics, such as handwriting and drawing, use specific drawing and handwriting tasks performed by the subjects involved in data collection [4]. From a health perspective, online handwriting biometrics are more appealing and informative on states of diseases, such as dementia, than other popular biometrics traits, such as fingerprints or iris [3], [4] because they make part of routine functional activities from which evidence are drawn affected by the disease.

In the last two decades, online handwriting processing has been employed in the computerized assessment of neurodegenerative disorders (e.g., Parkinson's disease (PD)) [5], [6]. Patients with PD experience progressive loss of dopaminergic neurons in substantia nigra pars compacta (located in the midbrain), which is consequently associated with cardinal motor symptoms such as bradykinesia, rigidity, resting tremor, or postural instability [7]–[9]. Therefore, especially during the clinical phase of the disorder, we can observe freezing of gait [10], hypokinetic dysarthria [11], hypomimia [12], or alterations in handwriting [6]. The latter was initially linked with micrographia, i.e., a progressive decline in amplitude (vertical micrographia) or with

The associate editor coordinating the review of this manuscript and approving it for publication was Gang Mei.

(horizontal micrographia) of letters [13]. Nevertheless, micrographia is one manifestation of altered handwriting in patients with PD. Others include more pronounced changes in kinematics and dynamics too. Letanneux et. reported a connection to developmental dysgraphia and proposed a new and more general term, PD dysgraphia [14].

Recently, several designs of decision support systems for diagnosing different PDs based on speech/voice analysis [15]–[20] or gait monitoring [21]–[23], have been proposed. However, compared with online handwriting processing, both speech assessment and gait monitoring require more technical equipment and are vulnerable to low signal quality due to a noncontrolled environment. Speech assessment requires high-quality recording conditions without background noise and further postprocessing of recorded speech. This includes human-operated speech segmentation, making the whole process more difficult. Gait monitoring or tremor assessment techniques require specialized equipment, such as motion capture systems, accelerometers, and gyroscopes. However, the diagnosis of PD using handwriting processing can be easily administered at the clinic or a patient's home. Handwriting acquisition is simple and natural and requires no timing or exhaustive repetitions.

A comprehensive review of quantitative analysis of PD dysgraphia and its computerized diagnosis for published works until 2019 has been summarized [4]–[6], [14]. Furthermore, we review the state-of-the-art designs published in 2020 and 2021, focusing on articles using online handwriting.

The rest of the paper is organized as follows: Section II reviews related works and presents state-of-the-art results obtained in PD diagnosis based on the PaHaW database. Section III describes the H2O platform used in this work. Section IV describes the PAHAW database. Section V describes the feature extraction process used and describes the type of feature obtained. Section VI presents a brief explanation of the modified version of the fast correlation-based filtering feature selection methodology. Section VII describes the front-end hyperparameters. Section VIII describes the experiments conducted and their results. Finally, in Section IX, we present final remarks, comments, and conclusions.

II. RELATED WORKS

This section reviews related works and state-of-the-art results obtained for PD diagnosis. Table 1 shows a summary of the state-of-the-art results.

Ammour *et al.* [24] quantitatively analyzed online handwriting in 28 PD patients and 28 age-matched healthy controls (HC). They quantified the performance of subjects (when writing a text in Arabic letters) according to 1482 kinematics (velocity, acceleration, jerk, etc.), dynamic (pressure, tilt, azimuth, etc.), temporal (e.g., duration), and some additional features. From a semi-supervised approach (employing cluster analysis), they differentiated the PD group with 71.44% accuracy. Furthermore, they concluded that the complications of fine motor skills in PD patients were mainly manifested in the kinematic feature set.

Liaqat Ali, *et al.* [25] propose a method for dealing with the highly unbalanced handPD dataset. To improve the PD detection accuracy on this dataset, they developed a cascaded learning system that cascades a Chi2 model with an adaptive boosting (Adaboost) model. Experimental results confirmed that the proposed cascaded system outperforms six similar cascaded systems using six state-of-the-art machine learning models, respectively.

Taleb *et al.* [26] introduced a PD diagnosis concept that uses convolutional neural networks (CNN) fed by spectrograms (calculated from various online handwriting/drawing tasks) and CNN bidirectional long-short-term memory networks (CNN-BLSTM) fed by raw time series. In the publicly available dataset called HandPDMultiMC, containing 21 PD and HC, respectively, a classification accuracy of approximately 97.62% was achieved by combining CNN-BLSTM models trained with jittering and synthetic data augmentation. They trained 204,060 parameters model for one day using an NVIDIA GTX 1080 GPU of 8 GB.

Gupta *et al.* [27] explored the effect of age and gender on the performance of classification models. Thus, they used the PaHaW database [28] containing 37 PD patients and 38 HC. The subjects performed seven tasks including a sentence or isolated words. The data were parametrized using kinematic, entropic, and energetic features and fed into age- and gender-dependent support vector machine (SVM) models. A distinct set of discriminative features was observed in each category (age vs. gender). The results showed an improved classification accuracy of a general model from 75.76% to 83.75% and 79.55% in a female and male set, respectively.

Aouraghe *et al.* [29] focused on the effect of progressing fatigue in PD dysgraphia. They enrolled 40 PD patients and HC, respectively, copying a multiline paragraph in Arabic letters. First, the paragraph was segmented into individual lines and then, each line processed separately using a set of temporal, kinematic, dynamic, spectral, entropy-based, and wavelet-based features. The feature space was modeled by k-nearest neighbor classifier (KNN), SVM and decision trees. An accuracy of 92.86% was obtained when processing the last line of the paragraph, i.e., the line where the fatigue is mostly accented.

Deharab *et al.* [30] introduced a novel online handwriting parameterization using dynamic writing traces warping (DWTW).DWTW was applied to kinematic patterns of handwriting and returned parameters linked with the similarity between normative and pathological time series. The features were modeled using SVM and were evaluated on the PaHaW dataset (29 PD and 32HC; all eight tasks including handwriting and drawing of Archimedean spiral), and an accuracy of 88.33% was achieved.

TABLE 1. State -of -the -art in PD diagnosis based onUSING the PaHaW database. Legend: SVM – support vector machine; RF – random forests; ET – extremely randomized trees; ADA – AdaBoost, TKEO – Teager-Kaiser energy operator; EMD – empirical mode decomposition; DWTW – dynamic writing traces warping; CGP – cartesian genetic programming; 1DCL – 1-dimensional convolutional layer; BiGRUs – bidirectional gated recurrent units; ACC – accuracy; SEN – sensitivity; SPE – specificity; PRE – precision; REC – recall, AUC – area under the ROC curve.

Reference	Tasks	Features	Feature selection	ML algorithms	Performance	Conclusion
Drotar et al. 2016 [28]	all	temporal (e.g. duration), kinematic (e.g. velocity, acceleration, jerk), pressure-based.	Mann-Whitney U test	SVM	ACC = 81.3%, SEN = 87.4%, SPE = 80.9%	This is the study introducing the PaHaW dataset along with the baseline classification results.
Mucha et al. 2018 [35]	continuous and/or repetitive task, such as Archimedean spiral	Fractional-order-derivatives		SGBoost	ACC=97.14%	Features based on Fractional-Order Derivatives improve PD severity assesment.
Impedovo 2019 [34]	all	temporal (e.g. duration), kinematic (e.g. velocity, acceleration, jerk), dynamic (e.g. azimuth, tilt, pressure), entropic (e.g. Shannon entropy), TKEO, EMD-based, sigma-lognormal features, parameters based on the Maxwell-Boltzmann distribution and the discrete Fourier transform	Mann-Whitney U test, Relief	SVM	ACC=93.79% ACC = 98.44%	The accuracy for all task is 93.79. Accuracy of 98.44% was obtained when combining 3 tasks (Archimedean spiral, grapheme "l", and a word).
Angelillo et al. 2019 [42]	all	temporal (e.g. duration), spatial (e.g. length of stroke), kinematic (e.g. velocity, acceleration, jerk), dynamic (e.g. pressure), entropic (e.g. Shannon entropy), EMD-based, other (e.g. number of strokes)	SVM ranking	SVM	ACC = 91.67%	The best performance was obtained when combining 3 tasks (grapheme "l", bigram "le" and a word).
Diaz et al. 2019 [43]	all	static handwriting processed by CNN and enhanced by velocity and in-air movement	SVM ranking	ensemble (SVM, RF, ET, ADA)	ACC = 86.67%, SEN = 89.17%, SPE = 80.83%	The authors proved that a dynamic enhance of static handwriting could provide better results than considering the online/offline handwriting separately.
Mucha et al. 2019 [44]	all	kinematic features based on fractional calculus, dynamic (e.g. pressure, tilt)	-	XGBoost	ACC = 80.60%, SEN = 79.41%, SPE = 80.56%	Features based on fractional calculus enable complex assessment of kinematic aspects of graphomotor difficulties.
Gupta et al. 2020 [27]	all, excluding Archimedean spiral	kinematic (e.g. velocity, acceleration, jerk), entropic (e.g. Shannon entropy), and energetic (e.g. signal-to-noise ratio)	Mann-Whitney U test, SVM ranking	SVM	general: ACC = 75.76%, PRE = 97.72%, REC = 81.02%; gender-specific: ACC = 83.75%, SEN = 94.40%, SPE = 85.07%	Gender- and age-specific models reached better performance.
Deharab et al. 2020 [30]	all	features based on DWTW	-	SVM	ACC = 88.33%, SEN = 86.43%, SPE = 89.50%	The best performance was based on sentence, utilising only two features.
Taleb, et.al. 2020 [26]	all	horizontal and vertical displacment time series		CNN-BLSTM	ACC=97.62	High performance and high computer complexity for training.
Parziale et al. 2021 [31]	all	temporal (e.g. duration), kinematic (e.g. velocity, acceleration, jerk), dynamic (e.g. pressure), other (e.g. features base on rising/falling edge)	-	CGP	ACC = 71.18%, SEN = 70.33%, SPE = 73.83%	The classification model provides a good trade-off between discrimination power and interpretability.
Diaz et al. 2021[33]	all	spatial (e.g. x/y coordination, displacement), kinematic (e.g. velocity, acceleration, jerk), dynamic (e.g. pressure, tilt)	-	1DCL, BiGRUs	AUC = 96.88%, SEN = 92.50%, SPE = 100.00%	The study confirms the effectiveness of the sequence learning paradigm for processing sequential handwriting data.

Parziale *et al.* [31] addressed a recurrent issue in most published works, clinical interpretability. More specifically, authors frequently use handcrafted features poorly linked to physiological processes and employ less

interpretable machine learning models (so-called black boxes).

Such systems are unacceptable for clinicians. Therefore, cartesian genetic programming (CGP) (which provides a tradeoff between performance and interpretability) was used in comparison with the more common classifiers. The proposed methodology was evaluated on two datasets, PaHaW (37 PD and 38 HC; all tasks were used) and NewHandPD (31 PD and 35 HC; subjects performed a spiral and a meander). Using conventional temporal, kinematic, and dynamic features, CGP produced more accurate results than white-box methods (reaching 71.18% in PaHaW and 80.39% in NewHandPD) and more interpretable than the black boxes.

Lamba *et al.* [32] analyzed basic temporal (e.g., duration) and kinematic (e.g., velocity, acceleration, jerk) measures in 62 PD patients and 15 HC (enrolled in the frame of the Irvine (UCI) Parkinson's disease spiral drawings dataset). Due to high imbalance, the synthetic minority-oversampling technique was employed to balance the cohort. Next, data were modeled by several machine learning models, e.g., SVM, AdaBoost, and XGBoost. Finally, a classification accuracy of 96.02% was reported for AdaBoost.

Diaz *et al.* [33] discussed processed time series of online handwriting (including velocity, acceleration, jerk, displacement, pressure, etc.) using one-dimensional convolutions and bidirectional gated recurrent units (BiGRUs). This end-to-end pipeline was applied to PaHaW (37 PD and 38 HC; all tasks were used) and NewHandPD (31 PD and 35 HC; all tasks were used). The method provided competitive results (96.25% accuracy in PaHaW and 94.44% in NewHandPD), thus confirming the effectiveness of the sequence learning paradigm for processing sequential handwriting data.

Impedovo *et al.* [34] investigate different velocity-based signal processing techniques to address PD assessment. He uses kinematic, energy, and cepstral features. The energy and cepstral features are similar to the ones used in this work, but they do **not** use filterbank, and they so not use the filterbank output to calculate the cepstral. An accuracy result of 93.7% for all tasks, and 98.44% when he selects the top three tasks was reported.

Mucha *et al.* [35] combine kinematic features with fractional-order derivatives and reported an accuracy of 97.14%, for the continuous and repetitive task, such as Archimedean spiral.

Finally, in [36] we proposed the use of spectral and cepstral features for emotion recognition. Here, we concatenated these features with very simple temporal features.

This study explores new approaches of online handwriting parameterization, augmentation, analysis, and modeling with a special focus on improved diagnostic accuracy. Furthermore, we explore the impact of newly proposed spectral and cepstral features on classification accuracy and improve the pipeline by adding data augmentation and modified fast correlation-based filtering feature selection method.

TABLE 2. List of machine learning models.

Classification Algorithms
Deep learning (Deep neural networks-DNN)
Distributed random forest (DRF)
Extremely randomized trees (XRT) models)
Generalized linear model (GLM)
Generalized additive model (GAM)
Gradient boosting machine (GBM)
Naïve bayes classifier (NBC)
Rulefit (RF)
Stacked ensembles (SE)
XGBoost (XGB)
Support vector machine (SVM)

III. DATA MODELING

For data modeling, we use autoML H2O [37], [38]. Automatic machine learning (AutoML) is the process of automating algorithm selection, feature generation, hyperparameter tuning, iterative modeling, and model assessment. It eases training and evaluation of machine learning models. AutoML includes many ML models, however, we limit the number of models to the ones shown in Table 1. Also, it ensembles the best models that outperform individual models. Furthermore, it uses the area under the ROC curve as the default ranking metric for binary classification. The configuration is such that individual models are tuned using a two-fold cross-validation set. AutoML automatically performs Bayesian hyperparameter optimization.

Since a default performance metric for each machine learning task is specified internally, the leaderboard is sorted by that metric.

In Table 2, ML models include stacked ensemble models. The stacked ensemble is an efficient ensemble method, such that the predictions, from machine learning algorithms, are used as inputs in a second layer learning algorithm. In the second layer, H20 ensembles all models, (StackedEnsemble_AllModels), and the best of family, (StackedEnsemble_BestOfFamily), including the best models of each kind in the final ensemble.

IV. PAHAW DATABASE

This study employed the Parkinson's disease handwriting database (PaHaW), containing 37 PD patients and 38 age- and gender-matched HC enrolled at the department of neurology, St. Anne's university hospital in Brno [28]. Besides age and gender, the PD group is described in terms of PD duration, unified Parkinson disease rating scale part V score, and levodopa equivalent daily dose.

All subjects have no history or presence of any psychiatric symptom or disease affecting the central nervous system, except for PD. The acquisition was performed when the patients were in their ON state, i.e., approximately one hour after taking levodopa.

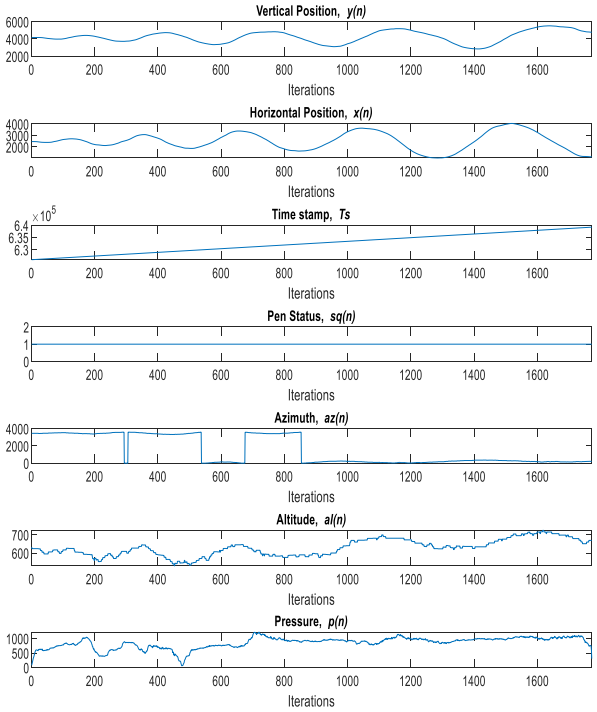

FIGURE 1. Online-drawing time series of the first task (Archimedean spiral drawings.

During the acquisition, the subjects were rested and seated in front of a table in a comfortable position. They completed a protocol on a printed template at a comfortable speed. The prefilled template was shown to the subjects; no restrictions on the number of repeated syllables/words in tasks or their heights were given. The signals were recorded at a 133 Hz sampling rate using the Intuos 4 M (Wacom technology) digitizing tablet and Wacom inking pen.

The protocol consists of the following eight tasks: Task 1 asks the user to draw, from inside out, an Archimedean spiral; tasks 2, 3, and 4 asks the user to repetitively write a cursive letter "l," syllable "le," and trigram "les," respectively; tasks 5, 6, and 7 asks the user to repetitively write a simple orthography and an easy syntax word, such that they are written in one continuous movement; finally, task 8 requires the user to write a longer sentence.

When the user was writing or drawing on the tablet (Fig. 1), the application captured the horizontal and vertical displacements of the pen tip in the x-axis, $x\,(n)$ and the y-axis, $y\,(n)$, respectively. Furthermore, the on-surface/in-air pen position information or status (touching/not-touching tablet's surface), $sq\,(n)$, the altitude of the pen with respect to the tablet's surface, $al\,(n)$, the pressure applied by the pen tip, $p\,(n)$, the azimuth angle of the pen with respect to the tablet's surface, $az\,(n)$, and the signal's timestamp, Ts, were captured.

A. DATA AUGMENTATION

Since the training database is small and unbalanced, we augment the smaller class such that both are equal in size. Then, we augment both classes to increase the training set.

Augmentation of the training data is performed as follows:

1. C_m = Identify the class with few samples.
2. N_s = Calculate the number of samples to compensate for the different number of samples.
3. Randomly select N_s of C_m.
4. For each selected sample, calculate the new feature vector by adding Gaussian random noise to the original features.

$$\mathrm{FV}_a = \mathrm{FV}_{u*} + \alpha * GV$$

where FV_{u*} is the feature vector of a random user, α is a value less than 0.2, and GV a vector with Gaussian random values. FV_a, FV_{u*}, and GV are vectors with equal dimensions.

V. FEATURE EXTRACTION

The front-end used in this study is shown in Fig. 2. This section describes the kinematic, statistics, spectral- and cepstral domain features used in the front-end. Definitions for calculating these features are provided in the next subsections and its graphical representation is shown in Fig. 3.

A. TEMPORAL AND KINEMATIC FEATURES

The row vector of temporal and kinematic features (KF) [34] for task τ and user u, applied to displacement, is defined as follows:

$$TKF^{\tau,u}_{d^{\tau,u}(n)} = \left[S_1^{\tau,u}, F_1^{\tau,u}, F_2^{\tau,u}, \dot{F}_1^{\tau,u}, \dot{F}_2^{\tau,u}, NCV^{\tau,u}, NCA^{\tau,u}, NCV_r^{\tau,u}, NCA_r^{\tau,u}, \curlywedge^{\tau,u}\right],$$

where,

$d^{\tau,u}(n) = \sqrt{x^{\tau,u}\,(n)^2 + y^{\tau,u}\,(n)^2}$, is the displacement,
$\mathcal{S}_1^{\tau,u}$ is the trajectory during handwriting divided by the duration of writing,
$F_1^{\tau,u} = \sum_{n=1}^{N-1} \mathscr{d}\,(2n)$, this is the on-air pen duration,
$F_2^{\tau,u} = \sum_{n=0}^{N-1} \mathscr{d}\,(2n+1)$, this is the on-paper pen duration,
$\mathscr{d}\,(i)$ is the duration of the stroke i; when $i \bmod 2 = 0$, the pen is on- air, otherwise it is on the tablet surface,
$\dot{F}_1^{\tau,u} = F_1^{\tau,u}/T$ represents the $F_1^{\tau,u}$ normalized to writing duration,
$\dot{F}_2^{\tau,u} = F_2^{\tau,u}/T$ represents the $F_2^{\tau,u}$ normalized to writing duration,
$\curlywedge^{\tau,u}$ is the ratio of time the pen spent in-air/on the tablet's surface,
$NCV^{\tau,u} = 1/\,(n-1)\sum_{i=1}^{N-1} |v\,(i) - v(i+1)|$ represents the number of changes in velocity direction (The mean number of local extrema of velocity),
$NCA^{\tau,u} = 1/\,(n-2)\sum_{i=1}^{N-2} |a\,(i) - a(i+1)|$ represents the number of changes in acceleration direction (The mean number of local extrema of acceleration),
$NCV_r^{\tau,u} = NCV^{\tau,u}/(T-1)$ represents the $NCV^{\tau,u}$ relative to writing duration,
$NCA_r^{\tau,u} = NCA^{\tau,u}/(T - 2 * Ts)$ represents the $NCA^{\tau,u}$ relative to writing duration, Ts is the sampling time

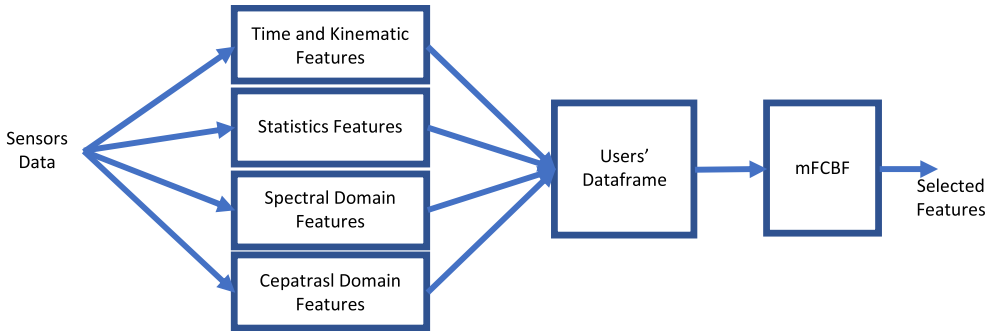

FIGURE 2. System Front-end mFCBF is the modified Fast Correlation-Based Filtering.

and $\mathcal{T}$ = *spiral*, *letterl*, syllable *le*, *trigrammles*, *word*1, *word*2, *word*3, *sentence* is the set of tasks to perform.

Then, the *TKF* row vector is the concatenation of *TKF* of each task τ; mathematically shown below, using relational algebra:

$$TKF^{U} = \left[\bigcup_{\tau \in \mathcal{T}} \left[TKF^{\tau,u}_{d(n)} \right]^{\mathrm{T}} \right]^{\mathrm{T}}.$$

We observe that to concatenate columns, we transpose the row vector, then we append the resulting column vectors using the union function. Finally, the row feature vector is obtained by transposing the column vector.

B. STATISTICS FEATURES

The statistics are obtained from the kinematic and stroke signals [39]. First, consider the set for task τ and user u:

$$g^{\tau,u}(n) = \left\{ k_w^{\tau,u}(n), \mathfrak{s}^{\tau,u}(n) \right\},$$

where

$n = 1, \ldots, T$,
$\mathfrak{s}^{\tau,u}(n)$ is the stroke signal,
$k_w^{\tau,u}(n) = \{v_w^{\tau,u}(n), a_w^{\tau,u}(n), j_w^{\tau,u}(n)\}$ is the set of kinematic signals, applied to signal in set $w^{\tau,u}(n)$,
$w^{\tau,u}(n) = \{d^{\tau,u}(n), x^{\tau,u}(n), y^{\tau,u}(n)$ is the set containing discrete, horizontal, and vertical displacements,
$d^{\tau,u}(n) = \sqrt{x^{\tau,u}(n)^2 + y^{\tau,u}(n)^2}$, is the discrete displacement,
$x^{\tau,u}(n)$ is the horizontal displacement,
$y^{\tau,u}(n)$ is the vertical displacement,
$v_w^{\tau,u}(n) = \frac{w^{\tau,u}(n) - w^{\tau,u}(n-1)}{T_s}$, is the velocity applied to signals in $w^{\tau,u}(n)$,
$a_w^{\tau,u}(n) = \frac{v_w^{\tau,u}(n) - v_w^{\tau,u}(n-1)}{T_s}$, is the acceleration applied to signals in $w^{\tau,u}(n)$,
$j_w^{\tau,u}(n) = \frac{a_w^{\tau,u}(n) - a_w^{\tau,u}(n-1)}{T_s}$, is the jerk applied to signal in $w^{\tau,u}(n)$, and $\mathcal{T}$ = *spiral*, *letterl*, syllable *le*, *trigrammles*, *word*1, *word*2, *word*3, *sentence* is the set of tasks performed for each user.

Statistics features row vector [Drotar *et al.*, 2014; 2016] for task τ and user u, is defined as follows:

$$SF_g^{\tau,u} = \left[\mathcal{B}_g^{\tau,u}, \mathfrak{M}_g^{\tau,u}, \mathcal{M}_g^{\tau,u} \right],$$

where

$\mathcal{B}^{\tau,u}_{g(n)}$ is the row vector of basic statistics features,
$\mathfrak{M}^{\tau,u}_{g(n)}$ is the row vector of mean features and
$\mathcal{M}^{\tau,u}_{g(n)}$ is the row vector of momentum features.
They are all applied to all signals in $g^{\tau,u}(n)$.
The row vector of basic statistics features is defined as follows:

$$\mathcal{B}_g^{\tau,u} = \left[\overset{\leftrightarrow}{g}^{\tau,u}, \breve{g}^{\tau,u}, \ddot{g}^{\tau,u}, \ddot{\bar{g}}^{\tau,u}, \dddot{\tilde{g}}^{\tau,u} \right],$$

where

$\overset{\leftrightarrow}{g}^{\tau,u}$ is the range,
$\breve{g}^{\tau,u}$ is the median,
$\ddot{g}^{\tau,u}$ is the mode,
$\ddot{\bar{g}}^{\tau,u} = \left(1/n \sum_{n=1}^{n=T} \left(\mathrm{g}^{\tau,u}(n) - \bar{g}^{\tau,u} \right) \right)^{1/2}$ is the standard deviation
$\dddot{\tilde{g}}^{\tau,u}$ is the outlier robust range (percentile 99th–percentile 1st); all above definitions applied to all signals in set $g^{\tau,u}(n)$, and
$\mathcal{T}$ is the set of tasks to perform.
The row vector of mean features is defined as follows:

$$\mathfrak{M}_g^{\tau,u} = \left[\bar{g}^{\tau,u}, \overline{\overline{g}}^{\tau,u}, \overbrace{g^{\tau,u}}^{tri} \right],$$

where

$\bar{g}^{\tau,u} = 1/n \sum_{n=1}^{n=T} g^{\tau,u}(n)$ is the arithmetic mean,
$\overline{\overline{g}}^{\tau,u} = \left(\prod_{t=1}^{t=T} g^{\tau,u}(n) \right)^{1/n}$ is the geometric mean,
$\overbrace{g^{\tau,u}}^{tri} = \bigcup \overbrace{g_i^{\tau,u}}^{tri} \forall i = 5, 10, 20, 30, 40, 50$, is the set of trimmed means for each of the values in i of $g^{\tau,u}(n)$; the trimmed mean is the mean after removing the outliers. For example, suppose $\tilde{g}_i^{\tau,u}$ has n values, the trimmed mean is the mean of $\tilde{g}_i^{C}$ excluding the highest and lowest k data values, where $k = n * (i/100)/2.$; all above definitions applies to all signals in $g^{\tau,u}(n)$, and
$\mathcal{T}(asdefined)$.

The row vector of momentum statistics features is defined as follows:

$$\mathcal{M}_g^{\tau,u} = \left[\overbrace{\mathcal{g}^{\tau,u}}^{qua} \overbrace{\mathcal{g}^{\tau,u}}^{per} \overbrace{\mathcal{g}^{\tau,u}}^{mom} \mathcal{k}^{\tau,u} \right],$$

where

$\overbrace{\mathcal{g}^{c}}$ *is* the row vector of quartiles ($Q_3 = 25(lower, Q_1 = 75/upper)$,

$\overbrace{\mathcal{g}^{\tau,u}}^{per} = \bigcup \overbrace{\mathcal{g}_i^{\tau,u}}^{per}, \forall i = 1, 5, 10, 20, 30, 90, 95, 100$, is the row vector of percentils,

$\overbrace{\mathcal{g}^{\tau,u}}^{mom} = \bigcup \overbrace{\mathcal{g}_i^{\tau,u}}^{mom}, \forall i = 3th, 4th, 5th, 6th,$ is the row vector of i moments,

$\mathcal{k}^{\tau,u} = 1/\sigma^4 \left(\sum_{n=1}^{n=T} \left(\mathcal{g}^{\tau,u}(n) - \bar{g}^{\tau,u} \right) \right)^{1/4}$, is the kurtosis; all above definitions applies to all signals in $\mathcal{g}^{\tau,u}(n)$, and $\mathcal{T} = \{spiral, letterl,$ syllable$le, trigrammles, word1, word2, word3, sentence$

Then, the row vector of the statistics feature for user U, using relational algebra, is shown below:

$$SF^U = \left[\bigcup_{\tau \in \mathcal{T}} \bigcup_{\mathcal{G} \in \mathcal{g}^C(n)} \left[SF_{\mathcal{G}}^{\tau,u} \right]^{\mathrm{T}} \right]^{\mathrm{T}}.$$

C. SPECTRAL-DOMAIN FEATURES

Spectral-domain feature row vectors for task τ and user u, applied to signals is $s^{\tau,u}(n)$, is defined as follows:

$$SDF_s^{\tau,u} = \left[FBCC_s^{\tau,u}(1), \ldots, FBCC_s^{\tau,u}(M) \right]$$

where

$LEFB_s^{\tau,u}(m) = \underset{\theta=1,2,\ldots,M}{filterbank} \{E_{\mathrm{s}}^{\tau,u}(k)\}$, *for* $m = 1, 2, \ldots, M$,

$E_{\mathrm{s}}^{\tau,u}(k) = log_2 \left(|S^{\tau,u}(k)|^2 \right)$, is the log energy spectrum,

$S^{\tau,u}(k) = \sum_{n=0}^{N-1} s^{\tau,u}(n) e^{-\frac{2\pi i}{N} kn}$, *for* $k = 0, 1, \ldots, K$, is the discrete Fourier transform of the signal and

$s^{\tau,u}(n) = x^{\tau,u}(n), y^{\tau,u}(n), p^{\tau,u}(n)$,

$x^{\tau,u}(n)$ is the horizontal displacement,

$y^{\tau,u}(n)$ is the vertical displacement, and

$p^{\tau,u}(n)$ is the pressure signal.

Then, the row vector of the spectrum-domain features is the concatenation of the *SDF* of each of the task τ for each signal in $s^{\tau,u}(n)$:

$$SDF^U = \left[\bigcup_{\tau \in \mathcal{T}} \bigcup_{\mathcal{S} \in \mathrm{s}} \left[SDF_{\mathcal{S}}^{\tau,u} \right]^{\mathrm{T}} \right]^{\mathrm{T}}.$$

D. CEPSTRAL DOMAIN FEATURES

Cepstral domain feature row vectors for task τ and user u, applied to signals in $s^{\tau,u}(n)$, is defined as follows:

$$CDF_s^{\tau,u} = \left[LEFB_s^{\tau,u}(1), \ldots, LEFB_s^{\tau,u}(M) \right]$$

where

$$FBCC_s^{\tau,u}(q) = \sum_{m=0}^{M-1} LEFB_s^{\tau,u}(m) e^{-\frac{2\pi \mathrm{i}}{N} qm},$$

$$For\ q = 1, 2, \ldots, Q.$$

M is the number of Filterbanks, $Q = M/2$ is the number of filterbanks, in the number of filterbanks divided by 2.

Then, the row vector of the cepstral domain features for user U, again using relational algebra, is shown below:

$$CDF^U = \left[\bigcup_{\tau \in \mathcal{T}} \bigcup_{\psi \in \Psi} \left[FBCC_s^{\tau,u}(q) \right]^{\mathrm{T}} \right]^{\mathrm{T}}.$$

E. USERS FEATURE

The row feature vector FV_u of each user, using relational algebra is shown as follows:

$$FV^{\mathbf{u}} = \left[\left[TKF^u \right]^T \cup \left[SF^u \right]^T \cup \left[SDF^u \right]^T \cup \left[CF^u \right]^T \right]^{\mathrm{T}}.$$

Alternatively, we define the disease state for each user as follows:

$$\mathcal{D}^u = \begin{cases} 0, \ Normal \\ 1, \ above\ Normal \end{cases} for\ all\ u = 1 \ldots U,$$

The row vector relating features to the emotional state is

$$FV\mathcal{D}^u = \left[\left[FV^{\mathbf{u}} \right]^{\mathrm{T}} \cup \left[\mathcal{D}^u \right]^{\mathrm{T}} \right]^{\mathrm{T}}$$

The data frame is defined as the union $FV\mathcal{D}^u$ of all users, and can be expressed using relational algebra notation as follows:

$$FV\mathcal{D} = \bigcup_{u=1}^{U} FV\mathcal{D}^u.$$

In this dataframe, the rows represent the number of users and the columns represent the features and its users' disease state.

VI. FEATURE SELECTION

Feature selection is a popular and common premodeling step in machine learning, especially in high-dimensional databases. Irrelevant features decrease the accuracy of data models because models also learn irrelevant information. Thus, selecting the right number of features increases the performance of the machine learning method.

Several methods for selecting features exist. All of them aim to obtain the best features and most do so by employing statistical tools with certain correlations to selection. The major difference between these tools is the selection criterion. Each patient has a considerable number of features, so we reduce the dimension of the number of features using a modified fast correlation-based filtering (FCBF) [40].

FCBF is based on two correlation factors: correlation between each feature and output and correlations among major difference between these tools is the selection criterion.

Each patient has a considerable number of features, so we reduce the dimension of the number of features using a modified fast correlation-based filtering (FCBF) [40].

FCBF selection is based on two steps. In the first step, the selected features are the ones whose correlation with the output are higher than a threshold. In the second step,

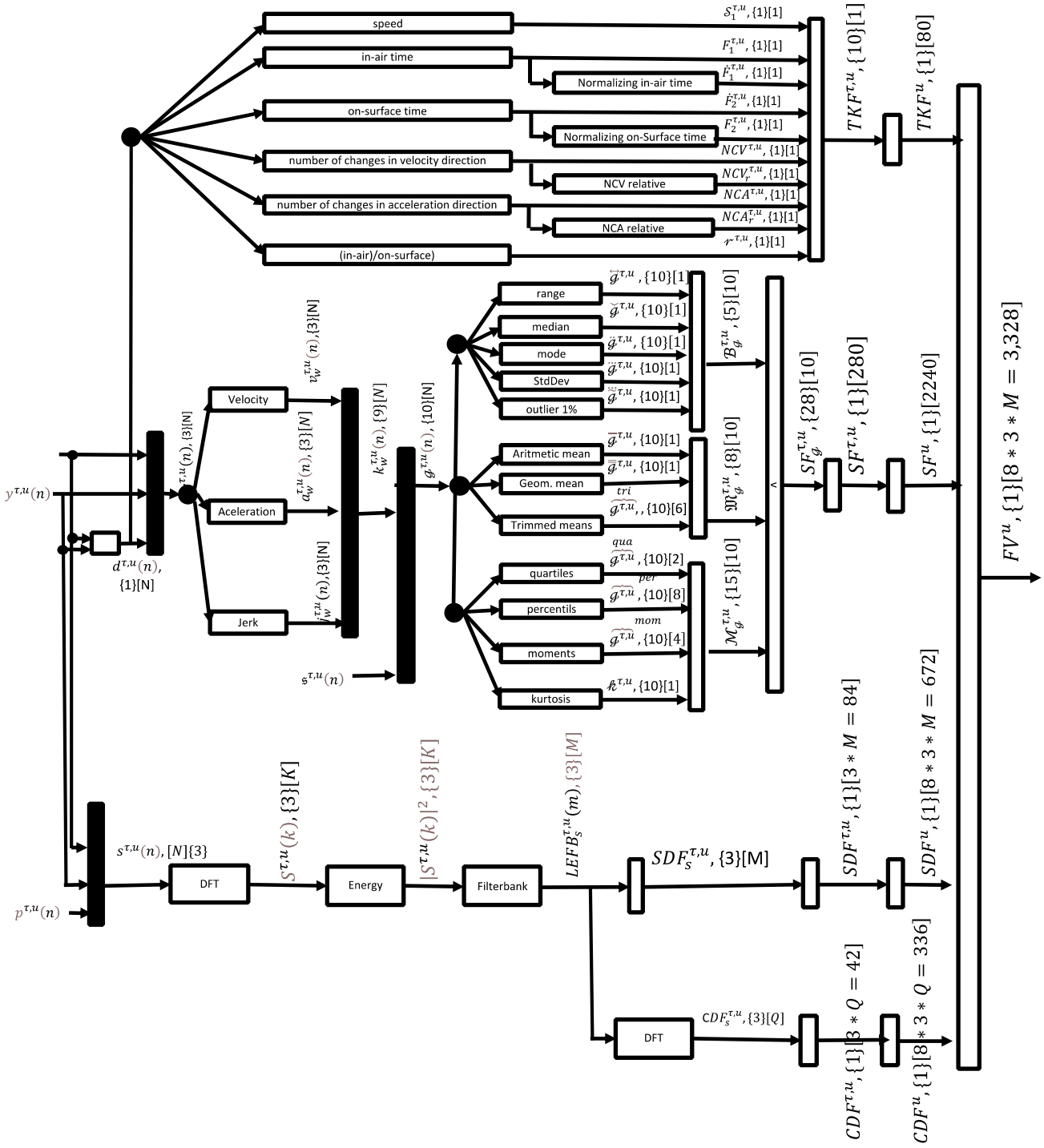

FIGURE 3. Feature Processing. Notation: {} indicates the number of members in the set; [] indicates the vector dimension of each element of the set. Black vertical rectangles indicate that the output is a set of the inputs; no-filled vertical rectangles indicate the outputs is a column concatenation of the inputs vectors. The processing of the horizontal rectangles is made for each element of the input set, creating an output of the same dimension as the input. In this figure, we assume a filterbanks dimension of 28 (M = 28) and cepstral coefficients number of 14 (Q = 14).

it takes the features of the first steps and selectes the ones with correlation lower than a threshold. In our modified version [36], mFCBF differs of the original FCBF at step 5, where the selected feature is the one having higher correlation with the output. Algorithm 1 shows a pseudo-code for the mFCBF process. mFCBF algorithm receives, as inputs, a dataframe and thresholds *oTh* and *iTh*. *oTh* is used to set the lower correlation value of each of the selected features and the output; *iTh* is used to set the higher correlation value between features. Using the values of *oTh* and *iTh*, we can

Algorithm 1 The mFCBF Algorithm Receives the Users Feature Matrix ($\mathbf{O^E}$), Minimum Correlation Threshold (**oTh**), and the Maximum Correlation Threshold (**iTh**) and Returns the Selected Set of Features

1: **function** mFCBF(**O**, **oTh**, **iTh**)
2: Calculate corr(**O**)
3: $\mathbf{O_{tmp}} \leftarrow$ Select columns whose output correlation is $>$ **oTh**
4: Calculate corr($\mathbf{O_{tmp}}$)
5: $\hat{O}_{\mathbf{Th,iTh}} \leftarrow$ Select columns whose correlation with the input is $<$ **iTh** and with the highest correlation with the output.
6: **return** ($\hat{O}_{\mathbf{Th,iTh}}$)
7: **end function**

find the right features to maximize the performance of the machine learning method. This operation is expressed as follows:

$$\widehat{FV\mathcal{D}}_{oTh,iTh} = mFCBF_{oTh,iTh}\left(FV\mathcal{D}\right).$$

Note that in $\widehat{FV\mathcal{D}}_{oTh,iTh}$ is a 2-D array, where one dimension represents the number of users and the other, the number of selected features.

Feature selectivity is controlled with *oTh* and *iTh* values. For example, given 370 user feature vectors, then for $iTh = 0.15$ and $oTh = 0.7$, the number of selected features reduce to 26, 28, and 20 for the depression, anxiety, stress states, respectively.

One way to visualize the relevance of features in improving performance is to use RadViz [41]. In RadViz, each data frame sample is represented inside the circle using the value in each series according to a physical metaphor. Each point is attached to each characteristic with a force proportional to the value the sample takes in the corresponding series. This implies that the final position is the equilibrium position between all forces representing the characteristics. Figs. 2 and 3 show the RadViz of the 2658 features and the selected 47 features, respectively. RadViz shows the dominant proportional values (DPV) of the features. In the graph, the higher cloud dispersion means higher DPV, whereas, a higher DPV means that features are easily exploited to improve the classification. Furthermore, these 47 features have higher DPVs than the complete set of 2658 features.

VII. FRONT-END HYPERPARAMETERS

Spectral-domain features (*SDF*) is a function of the filterbank bandwidth (*fbbw*), the bandwidth of the filters on the filterbank (*fbw*), the filterbank's initial frequency (*if*), and the overlap between filters on the filterbank (*ov*).

Conversely, features selection (FS) depends on the feature-output-correlation threshold (*oTh*) and the intra-feature correlation threshold (*iTh*).

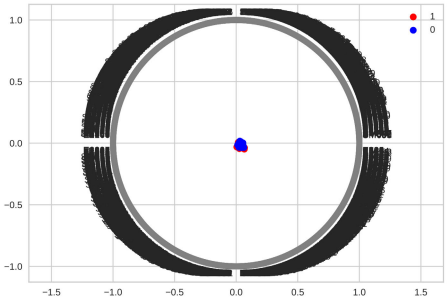

FIGURE 4. RadViz of the 2658 features. We can observe that there are no features with dominant proportional values (DPVs).

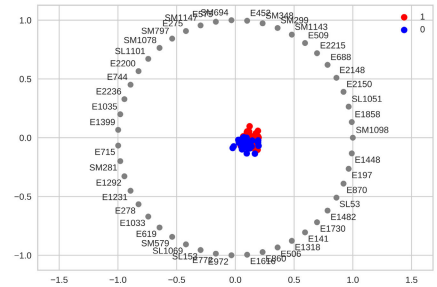

FIGURE 5. RadViz of 47 selected features. Note that selecting features increases the number of features with dominant proportional values (DPVs).

Therefore, the parameters for the final vector of features are

$$(fbbw, fbw, if, ov, oTh, iTh)$$

For practical, the range of values for each parameter is defined as follows:

$$\begin{aligned} iTh_{range} &= [0.2 - 1], \\ oTh_{range} &= [0 - 0.20], \\ fbbw_{range} &= [075], \text{ in Hz}, \\ fbw_{range} &= [0.5 - 3], \text{ in Hz}; \\ if_{range} &= [0.5] \text{ and} \\ ov_{range} &= [0] \text{ in \%}. \end{aligned}$$

A different set of features is selected for each combination of values. More so, each set of features produces a corresponding performance accuracy. Since one of these combinations is optimal, we find the combination that maximizes the ML accuracy.

Since augmentation of the training data, user selection, and Gaussian noise is random, we are unaware of which users and random sequences generate a better model. Therefore, we train and test the model for different sets of users and different random sequences, and we select the maximum accuracy.

VIII. RESULTS

The Leave-Percentage-Out (LPO) was used for testing. Here, the data model is trained with all database registers but a percentage, and the test is performed on registers that were out.

TABLE 3. Accuracy (%) results for different sets of coefficients with(FS)/without(NO_FS) feature selection applying mFCBF as features selection.

Features	NO_FS	FS
KF	80	88.57
$KF \cup SF$	80	94.28
$KF \cup SF \cup SDF$	82.85	97.14
$\boldsymbol{KF \cup SF \cup SDF \cup CDF}$	88.57	**98.57**

This was repeated until all possibilities were checked, then, we averaged the accuracy of all tests. In our experiments, we leave the 15% out.

We sample different filterbank's hyperparameters as follows:

$$\begin{aligned} fbbw^{s}_{range} &= [15, 20, \ldots 50], \text{ in Hz}, \\ fbw^{s}_{range} &= [0.2, 0.3, , \ldots, 1.0], \\ fi_{range} &= [0.5] \text{ and} \\ ov_{range} &= [0]. \end{aligned}$$

and, the mFCBF hyperparameters (oTh and iTh) are given as follows:

$$\begin{aligned} oTh^{s}_{range} &= [0, 0.02, 0.04, 0.06, \ldots, 0.18, 0.20], \\ iTh^{s}_{range} &= [0.2, 0.30, \ldots, 1.0], \end{aligned}$$

Therefore, we find the combination of this sample space that maximizes ML accuracy.

Table 3 shows the different accuracies for different feature sets. The accuracy results for TKF, when using feature selection or not, are 88.87% and 80%, respectively. The accuracy results for concatenating SF and using either feature selection or not are 94.28% and 80%, respectively.

From these two experiments, we find that adding statistics feature when combined with TKF improves the result accuracy.

Table 3 shows that the accuracy of the results when concatenating SDF using either feature selection or not are 97.14% and 82.85%, respectively. The last accuracy result is for concatenating CDF using either feature selection or not, are 98.57% and 88.57%, respectively.

The training data for all experiments were augmented by 80%, and the amplitude of the random Gaussian (α) was set to 0.2.

IX. CONCLUSION AND FUTURE WORK

We applied spectral and cepstral features on Parkinson's disease detection. Although spectral and cepstral features have been successfully applied for emotion detection, here, we concatenate them with kinetic and statistical features.

Similar features were used in [52] without the filterbank, thereby providing the flexibility for changing filterbank's bandwidth, filterbank's filters bandwidth, filterbank's filters overlapping, and filterbank's initial frequency to improve performance.

As the first step, we calculated TKF using the displacement signals; SF using displacement, and horizontal and vertical displacement; the SDF and CDF from the displacement and the horizontal and vertical displacement, and pressure signals.

Since the employed dataset (PaHaW) contains 37 PD patients and 38 HC subjects, then as a second step, we augmented the smaller class of the training set such that both are equal in size; next, we augment both classes of the training data by randomly selecting 80% of the training patient's data and added random Gaussian noise in all augmentations. For the third step, we selected the most relevant features using mFCBF method. Finally, autoML was employed to train and test more than ten plain and ensembled classifiers.

Experimental results show that adding spectral and cepstral features to the kinematics and statistics features highly improves the classification accuracy, reaching a combined classification accuracy of 98.57%. This result shows that our proposed model outperforms the best state-of-the-art result, which sits at 97.62%. Moreover, the state-of-the-art model has higher computational complexity and is required to train 204,060 parameters model for one day using an NVIDIA GTX 1080 GPU of 8 GB.

REFERENCES

[1] M. Faundez-Zanuy, J. Fierrez, M. A. Ferrer, and R. Plamondon, "Handwriting biometrics: Applications and future trends in e-security and e-health," *Cogn. Comput.*, vol. 12, pp. 940–953, Aug. 2020.

[2] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, Jan. 2000.

[3] S. Z. Li and A. K. Jain, *Encyclopedia of Biometrics*. New York, NY, USA: Springer, 2015.

[4] C. De Stefano, F. Fontanella, D. Impedovo, G. Pirlo, and A. S. di Freca, "Handwriting analysis to support neurodegenerative diseases diagnosis: A review," *Pattern Recognit. Lett.*, vol. 121, pp. 37–45, Apr. 2018.

[5] D. Impedovo and G. Pirlo, "Dynamic handwriting analysis for the assessment of neurodegenerative diseases: A pattern recognition perspective," *IEEE Rev. Biomed. Eng.*, vol. 12, pp. 209–220, 2018.

[6] G. Vessio, "Dynamic handwriting analysis for neurodegenerative disease assessment: A literary review," *Appl. Sci.*, vol. 9, no. 21, p. 4666, Nov. 2019.

[7] W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brundin, J. Volkmann, A. E. Schrag, and A. E. Lang, "Parkinson disease," *Nat. Rev. Dis. Prim.*, vol. 3, no. 17013, pp. 1–21, 2017.

[8] R. B. Postuma, D. Berg, M. Stern, W. Poewe, C. W. Olanow, W. Oertel, J. Obeso, K. Marek, I. Litvan, A. E. Lang, G. Halliday, C. G. Goetz, T. Gasser, B. Dubois, P. Chan, B. R. Bloem, C. H. Adler, and G. Deuschl, "MDS clinical diagnostic criteria for Parkinson's disease," *Movement Disorders Off. J. Movement Disorder Soc.*, vol. 30, no. 12, pp. 1591–1601, Oct. 2015.

[9] J. Stochl, A. Boomsma, E. Ruzicka, H. Brozova, and P. Blahus, "On the structure of motor symptoms of Parkinson's disease," *Movement Disorders*, vol. 23, no. 9, pp. 1307–1312, 2008.

[10] J. Jankovic, "Parkinson's disease: Clinical features and diagnosis," *J. Neurol., Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008.

[11] J. Rusz, T. Tykalova, L. O. Ramig, and E. Tripoliti, "Guidelines for speech recording and acoustic analyses in dysarthrias of movement disorders," *Movement Disorders*, vol. 36, no. 4, pp. 803–814, Apr. 2021.

[12] S. Argaud, M. Vérin, P. Sauleau, and D. Grandjean, "Facial emotion recognition in Parkinson's disease: A review and new hypotheses," *Movement Disorders*, vol. 33, no. 4, pp. 554–567, Apr. 2018.

[13] H.-I. Ma, W.-J. Hwang, S.-H. Chang, and T.-Y. Wang, "Progressive micrographia shown in horizontal, but not vertical, writing in Parkinson's disease," *Behavioural Neurol.*, vol. 27, no. 2, pp. 169–174, 2013.

[14] A. Letanneux, J. Danna, J.-L. Velay, F. Viallet, and S. Pinto, "From micrographia to Parkinson's disease dysgraphia," *Movement Disorders*, vol. 29, no. 12, pp. 1467–1475, Oct. 2014.
[15] L. Brabenec, J. Mekyska, Z. Galaz, and I. Rektorova, "Speech disorders in Parkinson's disease: Early diagnostics and effects of medication and brain stimulation," *J. Neural Transmiss.*, vol. 124, no. 3, pp. 303–334, Mar. 2017.
[16] L. Moro-Velazquez and N. Dehak, "A review of the use of prosodic aspects of speech for the automatic detection and assessment of Parkinson's disease," in *Automatic Assessment of Parkinsonian Speech* (Communications in Computer and Information Science). Cham, Switzerland: Springer, Sep. 2020, pp. 42–59.
[17] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102418.
[18] L. Ali, C. Zhu, M. Zhou, and Y. Liu, "Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection," *Expert Syst. Appl.*, vol. 137, pp. 22–28, Dec. 2019.
[19] B. E. Sakar, M. E. Isenkul, C. O. Sakar, A. Sertbas, F. Gurgen, S. Delil, H. Apaydin, and O. Kursun, "Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 4, pp. 828–834, Jul. 2013.
[20] L. Ali, C. Zhu, Z. Zhang, and Y. Liu, "Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network," *IEEE J. Transl. Eng. Health Med.*, vol. 7, pp. 1–10, 2019.
[21] E. Balaji, D. Brindha, and R. Balakrishnan, "Supervised machine learning based gait classification system for early detection and stage classification of Parkinson's disease," *Appl. Soft Comput.*, vol. 94, Sep. 2020, Art. no. 106494.
[22] D. Joshi, A. Khajuria, and P. Joshi, "An automatic non-invasive method for Parkinson's disease classification," *Comput. Methods Programs Biomed.*, vol. 145, pp. 135–145, Jul. 2017.
[23] L. Brognara, P. Palumbo, B. Grimm, and L. Palmerini, "Assessing gait in Parkinson's disease using wearable motion sensors: A systematic review," *Diseases*, vol. 7, no. 1, p. 18, Feb. 2019.
[24] A. Ammour, I. Aouraghe, G. Khaissidi, M. Mrabti, G. Aboulem, and F. Belahsen, "A new semi-supervised approach for characterizing the Arabic on-line handwriting of Parkinson's disease patients," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 104979.
[25] L. Ali, C. Zhu, N. A. Golilarz, A. Javeed, M. Zhou, and Y. Liu, "Reliable Parkinson's disease detection by analyzing handwritten drawings: Construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *IEEE Access*, vol. 7, pp. 116480–116489, 2019.
[26] C. Taleb, L. Likforman-Sulem, C. Mokbel, and M. Khachab, "Detection of Parkinson's disease from handwriting using deep learning: A comparative study," *Evol. Intell.*, pp. 1–12, Sep. 2020.
[27] U. Gupta, H. Bansal, and D. Joshi, "An improved sex-specific and age-dependent classification model for Parkinson's diagnosis using handwriting measurement," *Comput. Methods Programs Biomed.*, vol. 189, Jun. 2020, Art. no. 105305.
[28] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease," *Artif. Intell. Med.*, vol. 67, pp. 39–46, Feb. 2016.
[29] I. Aouraghe, A. Alae, K. Ghizlane, M. Mrabti, G. Aboulem, and B. Faouzi, "A novel approach combining temporal and spectral features of Arabic online handwriting for Parkinson's disease prediction," *J. Neurosci. Methods*, vol. 339, Jun. 2020, Art. no. 108727.
[30] E. D. Deharab and P. Ghaderyan, "Parkinson's disease detection using dynamic writing traces warping," in *Proc. 28th Iranian Conf. Electr. Eng. (ICEE)*, Aug. 2020, pp. 1–4.
[31] A. Parziale, R. Senatore, A. Della Cioppa, and A. Marcelli, "Cartesian genetic programming for diagnosis of Parkinson disease through handwriting analysis: Performance vs. interpretability issues," *Artif. Intell. Med.*, vol. 111, Jan. 2021, Art. no. 101984.
[32] R. Lamba, T. Gulati, K. A. Al-Dhlan, and A. Jain, "A systematic approach to diagnose Parkinson's disease through kinematic features extracted from handwritten drawings," *J. Reliable Intell. Environ.*, vol. 7, no. 3, pp. 253–262, Sep. 2021.
[33] M. Diaz, M. Moetesum, I. Siddiqi, and G. Vessio, "Sequence-based dynamic handwriting analysis for Parkinson's disease detection with one-dimensional convolutions and BiGRUs," *Expert Syst. Appl.*, vol. 168, Apr. 2021, Art. no. 114405.
[34] D. Impedovo, "Velocity-based signal features for the assessment of Parkinsonian handwriting," *IEEE Signal Process. Lett.*, vol. 26, no. 4, pp. 632–636, Apr. 2019.
[35] J. Mucha, J. Mekyska, Z. Galaz, M. Faundez-Zanuy, K. Lopez-de-Ipina, V. Zvoncak, T. Kiska, Z. Smekal, L. Brabenec, and I. Rektorova, "Identification and monitoring of Parkinson's disease dysgraphia based on fractional-order derivatives of online handwriting," *Appl. Sci.*, vol. 8, no. 12, p. 2566, Dec. 2018.
[36] J. A. Nolazco-Flores, M. Faundez-Zanuy, O. A. Velázquez-Flores, G. Cordasco, and A. Esposito, "Emotional state recognition performance improvement on a handwriting and drawing task," *IEEE Access*, vol. 9, pp. 28496–28504, 2021.
[37] *AutoML in H2O.Ai*. H2O.Ai. Accessed: Aug. 11, 2021. [Online]. Available: https://www.h2o.ai/products/h2o-automl/
[38] *The H2O Python Module*. The H2O Python Module—H2O Documentation. Accessed: Aug. 11, 2021. [Online]. Available: https://docs.h2o.ai/h2o/latest-stable/h2o-py/docs/intro.html
[39] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Analysis of in-air movement in handwriting: A novel marker for Parkinson's disease," *Comput. Methods Programs Biomed.*, vol. 117, no. 3, pp. 405–411, 2014.
[40] L. Yu and H. Liu, "Feature selection for high-dimensional data: A fast correlation-based filter solution," in *Proc. 12th Int. Conf. Mach. Learn. (IGML)*, 2003, pp. 856–863.
[41] *RadViz Visualizer*. RadViz Visualizer—Yellowbrick v1.3.Post1 Documentation. Accessed: Aug. 11, 2021. [Online]. Available: https://www.scikit-yb.org/en/latest/api/features/radviz.html
[42] M. T. Angelillo, D. Impedovo, G. Pirlo, and G. Vessio, "Performance-driven handwriting task selection for Parkinson's disease classification," in *AI*IA 2019—Advances in Artificial Intelligence* (Lecture Notes in Computer Science), 2019, pp. 281–293.
[43] M. Diaz, M. A. Ferrer, D. Impedovo, G. Pirlo, and G. Vessio, "Dynamically enhanced static handwriting representation for Parkinson's disease detection," *Pattern Recognit. Lett.*, vol. 128, pp. 204–210, Dec. 2019.
[44] J. Mucha, M. Faundez-Zanuy, J. Mekyska, V. Zvoncak, Z. Galaz, T. Kiska, Z. Smekal, L. Brabenec, I. Rektorova, and K. Lopez-de-Ipina, "Analysis of Parkinson's disease dysgraphia based on optimized fractional order derivative features," in *Proc. 27th Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2019, pp. 1–5.

JUAN A. NOLAZCO-FLORES (Senior Member, IEEE) was born in Gral. Terán, Nuevo León, Mexico. He received the B.Sc. and M.Sc. degrees from the Tecnológico de Monterrey, Monterrey, México, in 1986 and 1988, respectively, and the M.Phil. and Ph.D. degrees from the University de Cambridge, Cambridge, U.K., in 1992 and 1995, respectively.

From 2017 to 2020, he was the Dean of the School of Engineering and Science, South Region, Tecnológico de Monterrey. From 2011 to 2017, he was the Head of Information Technology, Tec of Monterrey. He was a Visiting Scholar with Marburg University and Mannheim University, Germany; Carnegie Mellon University, USA; Zaragoza University, Spain; and Cambridge University, U.K. He is currently the Head of the Data Science Hub, Tec of Monterrey. He is the author of more than 60 articles indexed in the Scopus citation report. His research interests include signal processing and data science applied to speech, security, and health.

Dr. Nolazco-Flores is a member of the Mexican Academy of Science, a member of the Mexican Research System, and an ACM Senior Member. He was awarded in 2005 and 2009 the ITESM Teaching and Research Award. He is a certified ABET Evaluator, and has participated as an Evaluator of the computer science programs at the University of West Florida; Grand Canyon University, Phoenix, NV, USA; and University of Iowa, Iowa City, IA, USA.

MARCOS FAUNDEZ-ZANUY was born in Barcelona, Spain. He received the Ph.D. degree from the Polytechnic University of Catalunya, in 1998.

From 2009 to 2018, he was the Dean of Escola Superior Politecnica Tecnocampus (the Polytechnic University of Catalunya, in 2015, and Pompeu Fabra University, after 2015. From 2010 to 2019, he was the Head of Research at Tecnocampus. He is currently a Full Professor at ESUP Tecnocampus Mataro, where he is the Head of the Signal Processing Group. He is the author of more than 100 articles indexed in ISI journal citation report, more than 100 conference papers, around ten books, and responsible for ten national and European research projects. His research interests include biometrics applied to security and health.

Dr. Faundez-Zanuy was the Initiator and the Chairperson of the European COST Action 277 Nonlinear Speech Processing, and the Secretary of COST Action 2102 Cross-Modal Analysis of Verbal and Nonverbal Communication. He is a Spanish Liaison of EURASIP.

V. M. DE LA CUEVA (Member, IEEE) was born in Xalapa, Veracruz, México. He received the B.Sc., M.Sc., and Ph.D. degrees from the Tecnológico de Monterrey, Monterrey, México, in 1986, 1993, and 1998 respectively.

From 1988 to 2021, he had been the Head of Department of Mechatronics on Campus Santa Fe; the Head of the School of Engineering; the Head of Postgraduate Studies of Engineering; and the Head of Research on Campus Ciudad de México, Tecnológico de Monterrey. He is currently the Head of the School of the Engineering and Science, Tecnológico de Monterrey, Toluca. He is the author of more than 30 articles published in journals and research congresses. He is the coauthor of three books in computer science and robotics. His research interests include artificial intelligence, robotics, and education.

JIRI MEKYSKA was born in Holesov, Czech Republic. He received the Ph.D. degree in telecommunications from the Brno University of Technology, Czech Republic, in 2014. He is currently the Head of the Brain Diseases Analysis Laboratory, Brno University of Technology, where he is also the Head of the Signal Processing Laboratory, Human–Machine Interaction Group, Department of Telecommunications. In cooperation with neurologists/psychologists from different countries, he develops diagnostic/monitoring systems focused on Lewy body diseases and developmental dysgraphia. He has coauthored more than 80 publications indexed by the Web of Science. He deals with the research of noninvasive neurodegenerative and neurodevelopmental disorders analysis based on speech, sleep, and handwriting processing.