A Novel Framework of Two Successive Feature Selection Levels Using Weight-Based Procedure for Voice-Loss Detection in Parkinson's Disease

AMIRA S. ASHOUR [1], **MAJID KAMAL A. NOUR** [2], **KEMAL POLAT** [3], **YANHUI GUO** [4], **WAFAA ALSAGGAF** [5], **AND AMIRA EL-ATTAR** [6]

[1]Electronics and Electrical Communications Engineering Department, Faculty of Engineering, Tanta University, Tanta 31527, Egypt
[2]Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[3]Department of Electrical and Electronics Engineering, Faculty of Engineering, Bolu Abant Izzet Baysal University, 14280 Bolu, Turkey
[4]Department of Computer Science, University of Illinois at Springfield, Springfield, IL 62703, USA
[5]Faculty of Computing and Information Technology, Information Technology Department, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[6]Department of Computer Science, Higher Institute of Management and Information Technology, Kafr El Sheikh 33516, Egypt

Corresponding author: Kemal Polat (kpolat@ibu.edu.tr)

ABSTRACT Parkinson's disease (PD) is one of the public neuro-degenerative disorders. Speech/voice disorder is considered one of the symptoms at an early stage. Acoustic and speech signal processing methods can potentially evaluate and measure PD-related vocal impairment. The present work proposed a novel feature selection framework using two levels of the feature selection procedure for voice-loss detection in PD patients. At the first level selection, the principal component analysis (PCA) and the eigenvector centrality feature selection (ECFS) methods are initially calculated independently, and the selected features from each method are considered as a separated sublist, namely ECFS selected features sublist, and PCA selected features sublist, in the first set. Accordingly, the first set, which is the first level selection set, is generated from the union of these two sublists using the top-selected features from both methods. In the training phase, a second level selection, which forms the second set (which is a subset from the first set), is generated to calculate the proposed weight of each selection method. Since in the present work, the ECFS provided superior performance to the PCA in the first level selection, the ECFS is applied to the first set in order to find weight values based on the contribution/ impact of the top-selected PCA- and ECFS- features in the second level. This weight is determined by finding a proposed ratio, which is multiplied directly by the selected ECFS features in the first level. The selected weighted ECFS features are then combined with the same PCA features to avoid ignoring any of the top-ranked features from the first level. This combination includes the final weighted-hybrid selected features that fed to a support vector machine (SVM) classifier to evaluate the proposed weighted hybrid selected features. Hence, in the test phase, the generated weight is used directly without any further need for the second level selection. Several comparative studies were conducted to evaluate the proposed feature selection performance for PD voice-loss detection. The experimental results established the superiority of the proposed procedure using cubic kernel-SVM with 94% accuracy for voice-loss detection in PD, while, with the same classifier, 88% accuracy was achieved without using the proposed selection method.

INDEX TERMS Parkinson's disease, voice loss, feature extraction, feature selection, principal component analysis, eigenvector centrality feature selection.

I. INTRODUCTION

Parkinson's disease is a progressive, enduring neurological disease leading to deterioration or death of the brain cells. It has several symptoms, including memory problems, as well

as depression, movement problems, including slowness, stiffness, and tremor. Also, walking/ balance problems, such as the freezing of the gait, are observed at the last stages of PD. Such indicators with their progression differ from patient to patient [1]. Generally, there are five main PD stages. In the earliest stage (stage 1), the PD patient suffers from mild symptoms on one side of the body, such as rigidity and tremor

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

in one hand or leg [2]. In stage 2, symptoms appear on the body's sides without balance impairment. In this stage, the PD patient suffers from speech abnormalities, loss of facial expression, the trunk's muscle stiffness, and stooped posture [3]. In the moderate stage (stage 3), the patient suffers from slow of movement, loss of balance, and occurrence of falling. In the last two severe stages (stages 4 and 5); patients are unable to live independently without assistance in walking and standing. Generally, in stage 5, the PD patient may fall while standing and may suffer from the freeze of gait while walking. Furthermore, the patient's physical and mental vitality decline [4]. Commonly, not all patients experience these five stages of PD progression in their order. In addition, the five stages vary in severity and time duration from patient to patient. In some cases, the PD patients facing all the five stages, whereas other patients skip from an early stage to an advanced one without suffering from the in-between stages. This leads to difficulty and complications in predicting PD progression, which attracts researchers to explore such an active research area and develop automated and accurate PD detection and prediction models [5]–[11].

For PD monitoring and telemedicine systems, early stages' detection and diagnosis become essential. In such cases, speech abnormalities are observed, including monotone voice, soft voice, slurring speech, and fading of the voice's volume after starting loud-speaking [12], [13]. These changes in the voice's characteristics in PD patients are considered the milestone for early detection of PD based on using features extraction procedures and classification techniques. Consequently, recent studies are concerned to use the linking between PD and speech loss and weakness, where, disordered vowels have widespread changes in the vibrations from approximately periodic to extremely aperiodic complex patterns. Such random properties in the PD patients' sounds facilitate the ability to extract an enormous number of dysphonia features. This leads to inconsistency in the feature space. Several diagnostic methods for Parkinson's disease were based on speech signals [14]–[23]. Moreover, different speech signal processing methods were designed for detecting PD cases from the variation in the speech signals and also grading the PD severity. For example, Sakar *et al*. [24] applied machine learning-based classification on huge, recorded voice samples of sentences, words, and sustained vowels in PD patients. The results established the discriminating characteristics in the sustained vowels compared to the other types of samples.

Generally, the association between various speech measures and movement (non-speech) measurements in PD patients has a great impact on the detection of PD cases. This assertion was verified by Goberman [25] through the analysis of the acoustic articulation, prosody, and phonation. The results depicted that seven acoustic measures out of sixteen speech features were considerably associated with movement measures, such as facial expression, gait, postural stability, posture, rest tremor, and postural tremor. These results established the correlation between the speech measures with both

axial- and non-axial- motor signs. Another study was conducted by Orozco-Arroyave *et al*. [26] on the recorded speech of different languages, namely German, Czech, and Spanish. Initially, a segmentation process was applied to the speech signals of utterances to separate the unvoiced and voiced frames. Afterward, unvoiced sounds' energy was modeled 25 bands using the Bark scale and 12 Mel-frequency Cepstral coefficients. This approach proved its accuracy to classify the speech signal of PD patients and healthy control individuals with an accuracy range of 85% to 99% based on the spoken language.

To separate healthy individuals from early-stage PD patients, Rusz *et al*. [27] assessed the impairment on the vocal to conclude the existence of the disorders in the speech at the PD early phases. The results proved that the main features to detect the PD cases are the dissimilarities in the fundamental frequency, where 78% of the early PD patients suffer from vocal impairment. Little *et al*. [28] extracted the fractal scaling and recurrence features during the speech analysis. These features were used to detect speech disorders as symptoms of the PD. Afterward; a bootstrapped classifier was applied to discriminate disordered from normal voices.

In a healthy voice, the pattern of vocal fold vibration has approximately periodic distribution. Thus, Tsanas *et al*. [29] extracted a huge number of extracted dysphonia features (132 dysphonia measures), including shimmer and jitter features as well as other features from the speech signals of both healthy and PD individuals. To overcome the uneven feature space with the possibility of over-fitting, different feature selection (FS) methods were compared, such as mRMR (minimum redundancy maximum relevance), LLBFS (local learning-based feature selection), LASSO (least absolute shrinkage and selection operator), and Relief [30]. Finally, support vector machines (SVMs) and random forest classifiers were used to classify to discriminate healthy individuals from PD cases. Recently, for PD patients' detection, Sakar *et al.* [16] extracted the tunable-Q wavelet coefficients and Mel-frequency Cepstral from the recorded voice signals of 252 individuals for feature extraction. Afterward, ensemble learning techniques were applied for classification after selecting the significant features based on their relevance using the mRMR FS scheme. The used ensembles entailed the k-nearest neighbor, multilayer perceptron, Random Forest, SVM with linear/ RBF kernels, logistic regression, and Naive Bayes classifier. The highest achieved accuracy was 86.0% with 0.84 F1-score by feeding the top 50 selected features to SVM with an RBF kernel classifier.

The preceding studies concluded the impact of the voice signal analysis to detect the PD cases, where several features can be extracted, such as glottis quotient, pitch-period entropy, F0-related measures, recurrence period density entropy, tunable-Q wavelet coefficients, Mel-frequency Cepstral, and empirical mode decomposition excitation ratio [16], [31]. This enormous number of different extracted features and dysphonia leads to inconsistent feature space with the possibility of over-fitting. This inspired the present work to

propose a novel FS procedure to reduce the feature subset and facilitate the interpretation of the speech/ voice signals to gain perceptions into the PD detection problem using the most significant features. This also guarantees the efficiency and power of the detection and classification model using a reduced number of features.

In the present work, the significant features in the extracted features, by Sakar *et al.* [16], are selected using the proposed novel feature selection framework for voice-loss detection of the PD patients. A first-level selection was proposed by selecting the significant features using both the principal component analysis (PCA) and the eigenvector centrality feature selection method (ECFS), separately, and then merging them by performing a union of the two subsets contain the top-ranked features of both. Afterward, the second level selection is conducted by applying other ECFS on the hybrid selected features to find a weight factor. Finally, the weighted-hybrid selected features were used as the feature vector space to be inputted to the cubic- SVM classifier for distinguishing the PD cases. Different comparative studies were also conducted to evaluate the performance of the proposed feature selection method for PD voice- loss detection.

The structure of the remaining sections is as follows. Section II includes the dataset description, background, and methodology of extracting features and applying feature selection methods by introducing the proposed procedure. In sections III and IV, the experimental results with comparative studies are reported and discussed, respectively. Finally, the proposed work is concluded in section V.

II. MATERIAL AND METHOD

A. DATASET

The used data set consists of voice signals recorded from 188 PD patients and 64 healthy persons as a control group. Each of these cases has 3 records or repetitions, which provided an overall number of samples equals 756 samples. Thus, the 753 features (attributes) were extracted from each patient in the dataset of 756 instances [16]. Speech features are employed to assess PD patients. The most popular speech baseline features are the fundamental frequency parameters, harmonicity parameters, jitter, shimmer, WT coefficients, MFCCs, recurrence period density entropy (RPDE), pitch period entropy (PPE), and Detrended fluctuation analysis (DFA) [16], [29]–[31]. In the present work, the used dataset entails 753 extracted features from speech signal processing procedures [16]. Such features include wavelet transform (WT) based features, time-frequency features, tunable Q-factor wavelet transform (TQWT) features, Mel frequency Cepstral coefficients (MFCCs), and vocal fold features which were extracted from the recorded speech of PD patients. These extracted features are related to the effects of the PD on the speech which include voice becomes softer, the speech might be incomprehensive, slurred, expressed rapidly, and the voice's tone might become monotone.

B. METHODOLOGY OF VOICE-LOSS DETECTION IN PARKINSON'S DISEASE

1) TRADITIONAL FEATURES EXTRACTION AND SELECTION

Several speech signal processing processes were applied to PD patients recorded speech for clinically extract convenient information for PD valuation. These features in [16] are used in the present work to evaluate the proposed FS method. The huge number of 753 features increases the feature space dimensionality and increase the existence of irrelevant features possibility. Accordingly, FS is the main process to reduce the dimensionality of the feature space by selecting significant features. FS techniques can be categorized into filter methods that use the proxy measure to score features, wrapper methods, and embedded methods [32]. Robust FS methods select significant features and discard irrelevant and redundant features [33]. This improves the data quality and the performance of the used classifier, accordingly. In the present work, the PCA and ECFS methods are applied as a hybrid combination procedure to gain the benefits of both of them.

a: PRINCIPAL COMPONENT ANALYSIS

The PCA is used to reduce the dimensionality of the feature space of the samples. This concept is achieved by transforming data into a new set of variables by calculating the eigenvectors and eigenvalues of the covariance matrix. In the PCA, the principal components are computed for an input matrix X of size $m \times n$, containing m samples of n features to find the eigenvalues and eigenvectors of the correlation matrix, which is given by:

$$\Sigma = Y^T Y \tag{1}$$

$$Y = X - \eta_X \tag{2}$$

where η_X is the mean value of the features. Hence, the principal components matrix is given by:

$$R = US \tag{3}$$

where S is a diagonal matrix of the singular values and U is an $n \times n$ matrix, also, the principal component R_j is given by:

$$R_j = s_j U_j \tag{4}$$

This component represents the scaled left-singular vector using the standard deviation of the data points in the consistent direction, where the data variance is given by:

$$\sigma = \sum_j \lambda_j \tag{5}$$

where λ_j is the eigenvalue. The output of the PCA is the principal component, where R contains the R_j principal components of the input samples. Therefore, in the present work, the PCA is applied to the feature space, and only the principal components of the input samples, which provided a good description of the data, were considered. Furthermore, the ECFS method is applied in the present work.

b: EIGENVECTOR CENTRALITY FEATURE SELECTION

The ECFS method orders the extracted features to rank the features according to their relevance. Eigenvector selection is a graph-based feature selection method for ranking features according to a graph centrality measure (Eigenvector centrality). It maps the features as nodes in a graph and scores the connected edges of the distributed features, where the edges define the path between features. For a set of nodes F which related to features $F = \{f_1, \ldots, f_n\}$, an undirected graph is defined as [34]:

$$G = (V; E) \qquad (6)$$

where V is the set of vertices corresponding to each feature f, and E represents the weighted edges between features. To define the nature of the weighted edges, an adjacency matrix A is represented, which is associated with G. In addition, each element a_{ij} of A, which represents a pairwise potential term, is given by:

$$a_{ij} = \varphi(f_i, f_j) \qquad (7)$$

where the potentials are represented as a binary function $\varphi(f_i, f_j)$ of the node. Consequently, the adjacency matrix of the graph can be expressed as [34]:

$$A = \alpha K + (1 - \alpha) \sum a_{ij} \qquad (8)$$

where A is the adjacency matrix of a directed graph is a matrix $n*n$, which contain nonnegative integers, such as $aij=$ number of arrows from all nodes. Also, α is a loading coefficient $\in [0, 1]$, and K is a kernel obtained using the Fisher criterion. Since the FS is used mainly to find discriminating features, only the features that can distinguish the classes are kept. Hence, mutual information is used to rank the features by assigning a high score to the features that highly can predict each class. This scoring can be expressed as follows:

$$S_i = \sum_{y \in Y} \sum_{Z \in F} p(z, y) \log(\frac{p(z, y)}{p(z)p(y)}) \qquad (9)$$

S_i is the scoring function used to select the features, Z is the feature set of features $F = \{f_1, \ldots, f_n\}$, and Y is the set of class labels. For a set of nodes, F is related to features $F = \{f_1, \ldots, f_n\}$, where $1 \leq i \leq n, n = 753$, and y represents the class labels. Also, $p(., .)$ is the joint probability distribution function of the features measured to calculates the likelihood of the two events (features and certain class) occurring together at the same time. Thus, the probability of the feature occurrence at the same time in a certain class is measured. Also, the score is computed to keep only the features that are related to or lead to these classes. In the ECFS method, an eigenvector A is calculated, which is defined as v_0, which is related to the largest eigenvalue representing the strength of the connection between the nodes (i.e., features).

2) PROPOSED WEIGHTED-BASED TWO-LEVEL FEATURE SELECTION HYBRIDIZATION

a: FIRST FEATURE SELECTION LEVEL (SET 1)

The proposed method consists of two-level selection, namely level 1, where both the PCA and the ECFS are used separately to obtain two separated feature vectors, PC_d and EV_b, including the selected principal components and the selected ECFS's features, respectively. Then, both PC_d and EV_b are combined to generate a new feature pool FV_1. For the m voice signal samples, the PC_d becomes a $m \times d$ matrix of d selected principal components and EV_b becomes a $m \times b$ matrix of b selected features using the ECFS. During this first level selection, the selected features in the PC_d were chosen as the top-ranked PCA features (PC) based on the superior obtained classification accuracy of the input samples in the training phase. The same is repeated during the selection of the best number of the selected features using the ECFS with the cubic-SVM to find the top-ranked ECFS features in EV_b. To gain the benefits of the two feature pools from PCA and ECFS, where these subsists considered two disjoint subists, the obtained selected features PC_d and EV_b are combined in the present work, leading to h selected features in FV_1, which can be expressed as follows:

$$FV_1 = [PC_d, EV_b] \qquad (10)$$

where FV_1 is the set of the combined feature vector of level 1 selection.

b: SECOND FEATURE SELECTION LEVEL (SET 2)

In the present work, the proposed second features selection level is applied to the FV_1 using the ECFS method again to determine the weights of the selected features. Consequently, the new generated feature vector FV_2 is expressed as:

$$FV_2 = [PC_k, EV_l] \qquad (11)$$

where PC_k and EV_l are the top-ranked selected features using the second level selection. Here, $k < d$ and $l \leq b$ as k and d are the number of the PC selected features in FV_1 and FV_2, respectively. In addition, l and b are the number of the selected EV features in FV_1 and FV_2, respectively. In addition, $g < h$, where g is the total number of selected features resulting from the second level selection including PC_k and EV_l. The set of the selected features from the second level (S_2) is considered a subset of the set of top- selected features from the first level selection (S_1). Consequently, a newly proposed non-zero weight w_{impact} is defined to indicate the ratio and impact of each selected feature from PCA features or the ECSF features that have the greatest contribution in FV_2, where the features from each method were labeled. This weight is proposed to increase the effect of the feature (impact) on the overall selected features set. This proposed weight factor is calculated using the following expression:

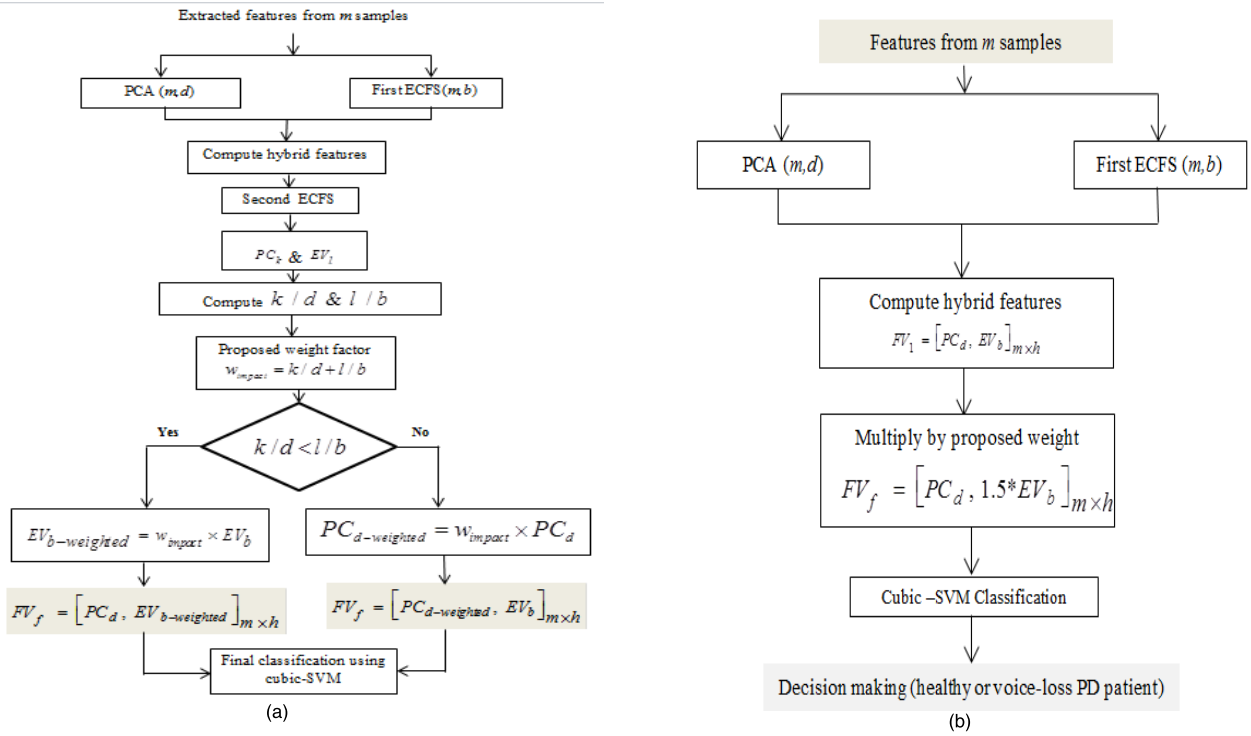$$w_{impact} = k/d + l/b \qquad (12)$$

FIGURE 1. Block diagram of the proposed system, where (a) training phase, and (b) test phase.

This proposed weight reflects the relation and contribution of the selected features using both methods. Thus, if $k/d > l/b$, then PC_d will be replaced by $w_{impact} \times PC_d$, while keeping EV_b without change. Otherwise, if $k/d < l/b$, then replace EV_b by $w_{impact} \times EV_b$ while keeping PC_d without change. In the present work, according to the reported results, l/b is greater than k/d, accordingly w_{impact} is multiplied by the selected features from the ECFS method in FV_1 using the following expression:

$$EV_{b-weighted} = w_{impact} \times EV_b \qquad (13)$$

Hence, the final hybrid combination consists of PC_d and $EV_{b-weighted}$, which can be given by:

$$FV_f = [PC_d, EV_{b-weighted}]_{m \times h} \qquad (14)$$

where FV_f includes h features in each sample, which are the modified selected features of first-level selection step using w_{impact}, which include the union of the selected features in the proposed method. Henceforth, the proposed weighted selected features assigned feature weights that signify the selected features prominence based on the achieved highest classification accuracy.

3) OVERALL PROPOSED VOICE-LOSS DETECTION IN PARKINSON'S DISEASE

After selecting the significant features and using the weight, the final pool of features FV_f is used in the classification-based PD detection. Typically, there are different types of machine learning classifiers that can be used to detect

voice-loss in PD. The weighted-hybrid features are fed to cubic-SVM. Also, a number of classifiers including the artificial neural network (ANN), and SVMs with different kernels were used for a comparative study. The proposed overall system of voice-loss detection is illustrated in Figure 1, which consists of two phases, namely train, and test. In the proposed feature selection method, the final, significant features in the selected features of both PCA and ECFS are weighted without ignoring any of these features. In the second level selection, the ECFS was applied for a second time to FV_1 for determining the value of w_{impact} based on the impact and the number of features in PC_k and EV_l. Then, the calculated weight factor is multiplied by EV_b, where the number of ECSF selected features in FV_2 is greater than the number of selected features from the PCA features, as proved in the results section, leading to the final selected feature subset FV_f. The final selected based weighting hybrid features are then used for final classification to detect the PD cases and distinguish the healthy and PD patient's classes. The FV_f is then inputted to the trained cubic-SVM classifier to attain the final PD detection using the classification results. These preceding steps are illustrated in Fig. 1 (a) in the training phase. In the training phase, the results of the second level selection depicted that the ECFS has a great impact compared to the PC selected features. Consequently, after calculating w_{impact} and based on the obtained results (in the Results section), w_{impact} it is multiplied by the ECFS selected features directly in the test phase. In this test phase, the combination of features FV_1 is used with the weight factor, which obtained from

the second level selection in the training phase, to determine the final selected feature vector FV_f. The final design of the proposed system is demonstrated in the test phase (Fig. 1(b)).

In Fig. 1, the final binary classifier using the selected features FV_f is able to accurately discriminate the PD patient cases from the healthy control ones.

III. RESULTS AND DISCUSSION

The proposed system was designed to reduce the dimensionality of the features with accurate classification for voice-loss detection in PD cases.

A. PERFORMANCE EVALUATION COMPARATIVE STUDY FOR DIFFERENT CLASSIFIERS WITHOUT FEATURE SELECTION

To ensure the efficiency of the proposed FS framework, a comparative study was conducted between the cubic-SVM and another 16 broadly used machine-learning procedures. The classification performance of several classifiers was evaluated using the feature subsets of 753 features from every 756 samples, as demonstrated in Figure 2.
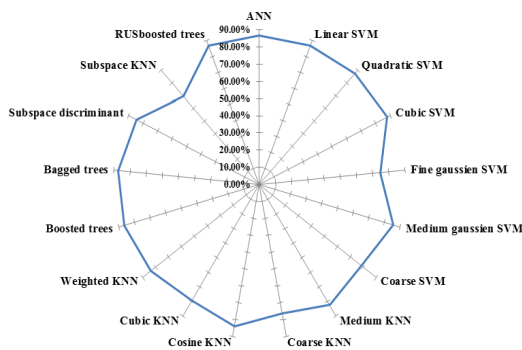
FIGURE 2. Comparative study of different classifiers accuracies in percentage using the 753 features.

Fig. 2 established the superiority of the SVM using a cubic kernel which achieved 88% accuracy. Thus, in the present work, the cubic- SVM has the best accuracy compared to using the other classifiers, including the ANN, Linear SVM, Quadratic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse SVM, Medium KNN, Coarse KNN, Cosine KNN, Cubic KNN, Weighted KNN, Boosted trees, Bagged trees, Subspace discriminant, Subspace KNN, and RUSboosted trees. Consequently, the cubic-SVM is used for the remaining cases in the current study of PD detection. In the present work, several sequential steps were performed and evaluated to reach the final proposed complete model evaluation as follows.

B. STEP 1: CUBIC- SVM BASED PCA

The performance of the cubic- SVM classifier with selecting the top-ranked principal components of the input samples using a different number of principle components r is reported in Table 1 as an example, where the other possibilities were also examined.

TABLE 1. Performance of cubic-SVM using PCA in the test phase.

Number of selected PC components (r)	Cubic- SVM Accuracy in %
753	**75.3%**
700	75.3%
500	81.9%
400	**83.9%**
300	83.9%
200	**91.1%**
100	91.1%
50	86.9%

Table 1 showed that the highest classification accuracy of value 91.1% is achieved by selecting the first 100 top-ranked principal components, which are used in the further coming steps as the selected features from the PCA.

C. STEP 2: CUBIC- SVM BASED ECFS

In this section, all the 753 extracted features from the voice signal in [16] were used. During the classification process, 70% of samples were used for training, while 30% were used for testing. Then, feature ranking using the ECFS method was applied to select the best top-ranked features. Table 2 reported the accurate measurements of the cubic-SVM classifier after features ranking using the first 753, 700, 600, 500, 300, 200, 100, and 50 ranked/ selected features for all samples, as an example of the obtained results from using grid search to find the best number of the selected ECFS features.

TABLE 2. Performance of cubic-SVM using ECFS in the test phase.

Number of selected ECFS features	Cubic- SVM Accuracy in %
753 (all features)	88.0
700	89.9
600	89.8
500	90.1
400	92.5
300	93.1
200	90.9
100	90.5
50	85.6

Table 2 depicted the best classification accuracy obtained with selecting the first 300 features using the ECFS method for voice- loss detection with 93.1% accuracy. Accordingly, we used these 300 ECFS top-ranked features in the further coming steps.

D. STEP 3: FIRST LEVEL SELECTION WITH THE GENERATED HYBRID SELECTED FEATURES (FV₁)

This step in the results section applied the procedure mentioned in section 2.1 of the first level selection. The proposed study gained the benefits of combining the selected features from the two FS methods in steps 1 and 2 to obtain the hybrid model of selected features. In the first level selection, 300 top-ranked selected features using ECFS(shown in step 1), and 100 top-ranked selected principal components of

PCA (in step 2) were integrated and summed up to a total of 400 features in a set called S1, i.e. $FV_1 = 400$, which are used in our proposed hybrid system. These 400 features were applied to a cubic-SVM, where the input has a size of 756×400, while the target was a 756×1 matrix with '1' indicating PD patient and '-1' indicating healthy individual, then the cubic-SVM achieved about 93.3% accuracy. This result indicated that the hybridization increases the classification performance compared to the preceding cases, namely using the whole extracted features without selection, using the ECFS only, and using the PCA only. This inspired the present work to study the impact of the selected ECFS and the selected PCA features in FV_1 using the proposed second level selection as follows, where the number of the selected features in the first level selection is found to be $N(FV_1) = 400$.

E. STEP 4: SECOND LEVEL SELECTION (FV_2) AND WEIGHT FACTOR IN THE TRAINING PHASE

This step in the results section applied the procedure mentioned in section 2.2 of the second level selection by calculating the weight factor in the training phase. To find the contribution/ impact of the different FS methods in FV_1, a second level selection was applied to the hybrid selected features $FV_1 = 400$ using another ECFS. Thus, the 400 combined and labeled selected features in the first feature selection are fed again to other ECFS to determine the top-ranked selected features for further use to determine the weight factor. Table 3 presented the accuracy of the cubic-SVM classifier after ranking using the first 200, 300, 350, and 370 ranked/ selected features for all samples from the 400 selected features as an example, where the classification accuracies were calculated for all ranking possibilities. Thus, Table 3 includes the worst and best cases.

TABLE 3. Performance of cubic-SVM using ECFS for second-level selection in the training phase.

Number of selected features in FV_2	Cubic- SVM Accuracy in %
200	92.6%
300	92.3%
350	93.8%
370	92.6%

Table 3 depicted that the highest classification accuracy of 93.8% is achieved by a new set called S2, using 350 top-ranked selected features from the hybrid selected features of ECFS and PCA. Since the selected features were labeled, it was found that the whole 300 ECSF features were included in this finally selected pool, while only 50 PCA features were used in this final selection stage out of a total 100 PCA features. Accordingly, the result showed that the percentage of using PCA is 50% (i.e., half of the PCA components that ranked in the second level selection) from the total number of selected features in the first level selection, while the percentage of using ECFS is 100% (i.e., all the ECSF that used and ranked in the first level selection were passed to

the second level selection). Accordingly, the weight factor was calculated using the obtained values of k, d, l and b to substitute in Eq. (12) as follows:

$$w_{impact} = k/d + l/b = (300/300) + (50/100) = 1.5 \quad (15)$$

Thus, in the present work, the weight factor is found to be $w_{impact} = 1.5$.

F. STEP 5: OVERALL PROPOSED FINAL SELECTED FEATURES SET (FV_f) AND VOICE LOSS DETECTION BASED CLASSIFICATION IN THE TEST PHASE

This step implements the test phase in the overall proposed framework shown in Fig. 1 (b) using the deduced weight value in the previous step. The calculated weight value $w_{impact} = 1.5$ is then multiplied by the selected ECFS features in FV_1 leading to $EV_{b-weighted}$ as given in Eq. (11), where the impact of the ECFS features is greater than the impact of the PCA features.

From the previous study, we found that the percentage of the selected features using ECFS is greater than PCA in the hybrid method, and the performance of selected features using eigenvector only is greater than PCA only. So, we multiplied the 300 selected features from eigenvector by the weight factor of 1.5 to increase the effect of the ECFS features on the overall hybrid selected features set in FV_1 to find the final weighted- hybrid selected features set FV_f. To prove the correctness of this way using the calculated weight factor by using (12), other weight values were tested and multiplied by the ECFS features and gathered with the PCA selected features with computing the value to find the relation between changing the weight and the classification accuracy as illustrated in Fig. 3 using FV_f in (14).
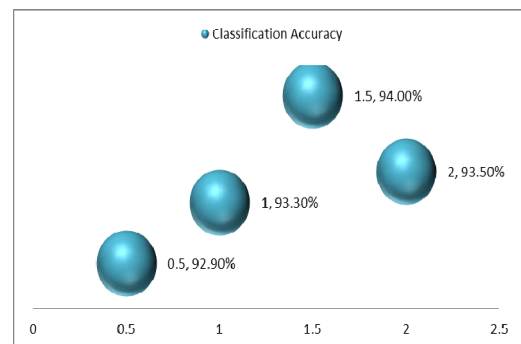
FIGURE 3. The relation between the change in the weight factor value and classification accuracy, where the x-axis indicates the weight factor value.

Fig.3 showed the performance of cubic-SVM using weighted hybrid features. It described the effect of changing the proposed weight value on the final classification accuracy. Accordingly, it is concluded that the proposed weight applied to the selected ECFS features improves the classification performance up-to 94% with a weight of value 1.5. This weight value equals the sum of the PCA contribution's percentage in the hybrid selected features set, and the ECFS selected features percentage in the second level ECFS. So, the

proposed weighted-hybrid selected features achieved the best accuracy to classify and detect the voice-loss in PD patients. Moreover, other classification performance metrics of the proposed weighted-hybrid selected features were measured. Such metrics include the sensitivity, specificity, precision, negative predictive value, miss rate (false negative rate), fall out (false positive rate), false discovery rate, false omission rate, and accuracy, which have the following values, 84.4%, 97.3%, 91.5%, 94.8%, 15.6%, 2.7%, 8.5%, 5.2%, and 94%, respectively. Finally, the Receiver Operating Characteristic (ROC) Curve, which is a probability curve for comparing diagnostic tests, is plotted in Fig. 4. Also, the area under the curve (AUC) indicates that the classification performance is close to the perfect classifier.
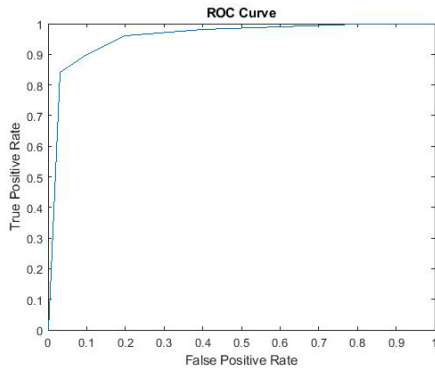
FIGURE 4. ROC of cubic SVM using the proposed weighted-hybrid selected features.

Figure 4 illustrated that the AUC equals 0.97 using the proposed method.

IV. DISCUSSION AND COMPARISON WITH STATE-OF-THE-ART WORK

In the proposed method, the weighted-hybridizing based second level selection (from training phase) was performed by merging both the PCA- based selected features and the weighted ECFS-based selected features. Lastly, the weighted-hybrid selected features were inputted to the SVM classifier to detect the voice- loss signals identifying PD cases. The preliminary feature selection using the PCA and the ECFS methods supported by the response of the cubic-SVM indicated the strength of features' association in the feature space. Nevertheless, ultimately, our purpose is to improve the association between the selected features using the PCA and ECFS. This goal was achieved by multiplying the selected features using the ECFS by a weight factor w_{impact}, where the accuracy of using EV_b separately is superior to using the PCA selected features PC_d. Consequently, the final binary classifier using the selected dysphonia measures FV_f was able accurately to discriminate the PD patient cases from the healthy control ones. To ensure the efficiency of the proposed method, a comparative study was conducted between the cubic-SVM and another 16 broadly used machine-learning procedures. This comparative study which reported in Fig. 2 proved the superiority of using the cubic-SVM. Furthermore, the pro-

posed method is superior by 8% improvement in the accuracy to the obtained results in [16] that used the same dataset, which achieved maximum accuracy of 86% with 0.84 F1-score and 0.59 MCC by feeding the top-50 features selected by mRMR to SVM-RBF classifier. Thus, the preceding results established the efficiency of the proposed selection method for the PD voice-loss detection based classification process. Subsequently, it is recommended to generalize this proposed new feature selection framework for automatic measurements and assessment method for PD patients at the early stage as well as in a range of different applied clinical purposes.

Moreover, a comparison between the proposed FS method on text clustering is compared to the hybrid feature selection method by Bharti and Singh in [35]. In [35], a modified union based on sets intersection was designed to avoid ignoring any of the selected features in the text clustering problem. The authors selected all top-ranked features from two sublists as well as the common features using the intersection between the non-selected features. By comparing this procedure with our proposed method, we represented the feature section levels in the methodology section in terms of the sets and sublists, where here the first set (S_1) consists of 400 features including the top-ranked features from first-level selection (300 selected features of ECFS and 100 selected features of PCA). In addition, the second set (S_2), which includes the second level top-ranked selected features, consists of 350 features (300 of ECFS and 50 of PCA). This relation between the two sets can be expressed as follows:

$$S_2 \subset S_1 \tag{16}$$

In terms of the union and intersection operations, the relation between the two sets, respectively, can be given as:

$$S_1 \cup S_2 = S_1 \tag{17}$$
$$S_1 \cap S_2 = S_2 \tag{18}$$

At the same time, the difference between the two proposed sets is given by:

$$S_1 - S_2 = S_3 \tag{19}$$

where S_3 consists of the features in S_1 and not exist in S_2, which includes the remaining 50 features. Since S_2 is a subset of S_1 without any remaining features in S_1 and there is no intersection between the remaining features sets, which is expressed as follows:

$$S_3 \cap \emptyset = \emptyset \tag{20}$$

Finally, by applying the concept of the modified union in [35] in our proposed method, the final set is can be formulated as:

$$S_F = (S_1 \cup S_2) \cup (S_3 \cap \emptyset) = S_1 \tag{21}$$

Accordingly, in the present work, the description of the two levels, where S_2 set is a subset of S_1 is illustrated in Figure 5.

According to the obtained results, the 400 features from the first level selection were used without ignoring any features and without any change, which achieved an accuracy
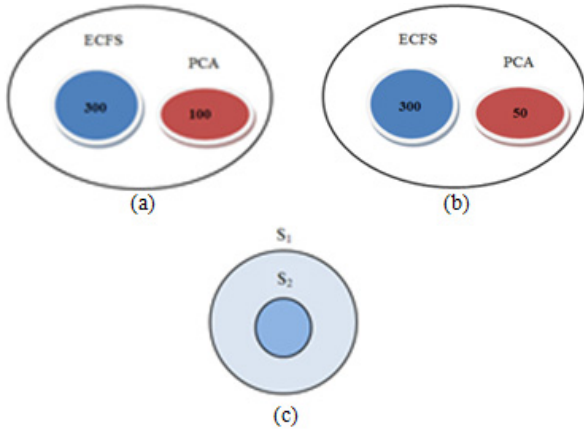
FIGURE 5. The relation between the proposed approach of the main two sets, where (a) represents set 1 (S_1) which includes the features of first-level selection method, (b) represents set 2 (S_2) includes the features of second-level selection method and (c) represents the union relation between the two sets.

of 93.3% classification- based PD detection. In addition, the probability of the PCA features existence in level two is:

$$P(PCA) = 50/100 = 0.5 \qquad (22)$$

Also, the probability of the existence of ECFS features in level two is:

$$P(ECFS) = 300/300 = 1 \qquad (23)$$

Subsequently, the proposed weighted-hybridizing based second level selection (from training phase) was performed by merging both the PCA-based selected features and the weighted ECFS-based selected features without ignoring any features. Lastly, the weighted-hybrid selected features were inputted to the SVM classifier to detect the voice-loss signals identifying PD cases, which achieved an accuracy of 94 %.

Generally, among the different feature selection methods that dealt with voice signals, both PCA and ECFS were recommended in several studies such as in [36]–[38]. Generally, PCA is very common as it projects the data into a new space with reducing the dimensionality of feature space. This guarantees the uncorrelated selected features from the PCA and any other feature selection method, where the selected PCA features are in a different space that differs from the space of any other features that can be selected using another feature selection method, such as the ECFS method. Hence, in the present work, the independently selected features from both the PCA and ECFS methods are uncorrelated, which improves the performance of the selection method. This independent and uncorrelated relation between the features from PCA and ECFS is also proved in the present work, where the intersection between their features is Ø. This accelerates the PD detection-based classification by getting rid of the correlated variables which do not contribute to any decision making.

Since the proposed method achieved significant results, it is recommended also to apply this proposed framework along with data discretization in data mining applications

compared to the Tsai, and Chen [39]. In addition, the proposed method can be incorporated with the feature selection methods in [40]–[44] as well as different FS methods that can be applied, such as random projection, independent component analysis (ICA), and non-negative factorization. However, due to the limitations of using tail-and-error to select the number of ranked/ selected features from each FS method in stages 1 and 2, it is recommended to automate this process using an optimization algorithm by using the classification accuracy as the fitness function. Also, it is recommended to generalize our efficient feature selection framework and test it with different problems and datasets, where it is based on the extracted features and their selection using our novel feature selection framework and did not depend directly on the nature of the signals and or the images (if used).

V. CONCLUSION
PD patients suffer from various symptoms in different body parts, including the tongue that leads to voice- loss. Recent researches are directed to find the association between speech impairment and PD for further use in PD cases detection and prediction. Dysphonia measures and other speech signal processing procedures are directed to expect the severity of PD symptoms using voice signals. Subsequently, numerous features can be extracted from the speech signal processing stage leading to over-fitting and the increased prospect of finding irrelevant features along with the problem of the imbalance in the datasets.

In this work, a novel cubic- SVM based weighted-hybrid two-level feature selection for voice-loss detection in PD was proposed. The first selection stage aims to select the significant features using both the ECFS and the PCA components for further hybridization. However, both FS methods do not achieve the same classification accuracy separately; this indicated that they do not have the same impact on the overall hybrid combination. Hence, the second level selection was proposed in the training phase only to find the proposed weight factor, which is multiplied by the hybrid selected features from the first selection of the most effective FS method. In the present work, the results proved that using the ECFS is superior to the PCA; accordingly, the final selected feature set included the same PCA selected features from the first selection stage while multiplying the selected ECFS features by 1.5 (computed weight value). Therefore, the proposed novel Cubic-SVM based two-level selection realized 94% classification accuracy, which is superior to several well-known machine learning classifiers. Overall, the proposed FS method established promising results in machine-learning based voice- loss detection in PD patients.

REFERENCES

[1] P. Calabresi, B. Picconi, L. Parnetti, and M. Di Filippo, "A convergent model for cognitive dysfunctions in Parkinson's disease: The critical dopamine–acetylcholine synaptic balance," *Lancet Neurol.*, vol. 5, no. 11, pp. 974–983, Nov. 2006.

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

[3] A. H. Friedlander, M. Mahler, K. M. Norman, and R. L. Ettinger, "Parkinson disease: Systemic and orofacial manifestations, medical and dental management," *J. Amer. Dental Assoc.*, vol. 140, no. 6, pp. 658–669, 2009.

[4] B. R. Bloem, Y. A. M. Grimbergen, M. Cramer, M. Willemsen, and A. H. Zwinderman, "Prospective assessment of falls in Parkinson's disease," *J. Neurol.*, vol. 248, no. 11, pp. 950–958, Nov. 2001.

[5] A. S. Ashour, A. El-Attar, N. Dey, M. M. A. El-Naby, and H. A. El-Kader, "Patient-dependent freezing of gait detection using signals from multi-accelerometer sensors in Parkinson's disease," in *Proc. 9th Cairo Int. Biomed. Eng. Conf. (CIBEC)*, Dec. 2018, pp. 171–174.

[6] A. El-Attar, A. S. Ashour, N. Dey, H. Abdelkader, M. M. A. El-Naby, and R. S. Sherratt, "Discrete wavelet transform-based freezing of gait detection in Parkinson's disease," *J. Exp. Theor. Artif. Intell.*, pp. 1–17, Sep. 2018.

[7] A. El-Attar, A. S. Ashour, N. Dey, H. Abdelkader, M. M. A. El-Naby, and F. Shi, "Hybrid DWT-FFT features for detecting freezing of gait in Parkinson's disease," in *Proc. ITITS*, 2018, pp. 117–126.

[8] H. L. Chen, C. C. Huang, X. G. Yu, X. Xu, X. Sun, G. Wang, and S. J. Wang, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," *Expert Syst. Appl.*, vol. 40, no. 1, pp. 263–271, 2013.

[9] M. Hariharan, K. Polat, and R. Sindhu, "A new hybrid intelligent system for accurate detection of Parkinson's disease," *Comput. Methods Programs Biomed.*, vol. 113, no. 3, pp. 904–913, Mar. 2014.

[10] M. Alhussein, "Monitoring Parkinson's disease in smart cities," *IEEE Access*, vol. 5, pp. 19835–19841, 2017.

[11] R. Das, "A comparison of multiple classification methods for diagnosis of parkinson disease," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1568–1572, Mar. 2010.

[12] M. Trail, C. Fox, L. O. Ramig, S. Sapir, J. Howard, and E. C. Lai, "Speech treatment for Parkinson's disease," *NeuroRehabilitation*, vol. 20, no. 3, pp. 205–221, 2005.

[13] C. Fox, L. Ramig, S. Sapir, B. Farley, A. Halpern, and J. Petska, "Voice and speech disorders in parkinson disease and their treatment," in *Neurorehabilitation in Parkinson's Disease: An Evidence Based Treatment Model* (Professional Book Division Thorofare). M. Trail, E. Protos, and E. Lai, Eds. Thorofare, NJ, USA: SLACK Inc.

[14] Y. Wu, P. Chen, Y. Yao, X. Ye, Y. Xiao, L. Liao, M. Wu, and J. Chen, "Dysphonic voice pattern analysis of patients in Parkinson's disease using minimum interclass probability risk feature selection and bagging ensemble learning methods," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–11, May 2017.

[15] Y. Wang, A.-N. Wang, Q. Ai, and H.-J. Sun, "An adaptive kernel-based weighted extreme learning machine approach for effective detection of Parkinson's disease," *Biomed. Signal Process. Control*, vol. 38, pp. 400–410, Sep. 2017.

[16] C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, M. Tutuncu, T. Aydin, M. E. Isenkul, and H. Apaydin, "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," *Appl. Soft Comput.*, vol. 74, pp. 255–263, Jan. 2019.

[17] C. R. Pereira, D. R. Pereira, S. A. Weber, C. Hook, V. H. C. de Albuquerque, and J. P. Papa, "A survey on computer-assisted Parkinson's disease diagnosis," *Artif. Intell. Med.*, vol. 95, pp. 48–63, Apr. 2019.

[18] S. A. Mostafa, A. Mustapha, M. A. Mohammed, R. I. Hamed, N. Arunkumar, M. K. Abd Ghani, M. M. Jaber, and S. H. Khaleefah, "Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease," *Cognit. Syst. Res.*, vol. 54, pp. 90–99, May 2019.

[19] S. Lahmiri and A. Shmuel, "Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine," *Biomed. Signal Process. Control*, vol. 49, pp. 427–433, Mar. 2019.

[20] S. Lahmiri, D. A. Dawson, and A. Shmuel, "Performance of machine learning methods in diagnosing Parkinson's disease based on dysphonia measures," *Biomed. Eng. Lett.*, vol. 8, no. 1, pp. 29–39, Feb. 2018.

[21] H. Gürüler, "A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method," *Neural Comput. Appl.*, vol. 28, no. 7, pp. 1657–1666, Jul. 2017.

[22] Z. Cai, J. Gu, C. Wen, D. Zhao, C. Huang, H. Huang, C. Tong, J. Li, and H. Chen, "An intelligent Parkinson's disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach," *Comput. Math. Methods Med.*, vol. 2018, pp. 1–24, Jun. 2018.

[23] D. Gupta, A. Julka, S. Jain, T. Aggarwal, A. Khanna, N. Arunkumar, and V. H. C. de Albuquerque, "Optimized cuttlefish algorithm for diagnosis of Parkinson's disease," *Cognit. Syst. Res.*, vol. 52, pp. 36–48, Dec. 2018.

[24] 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 Informat.*, vol. 17, no. 4, pp. 828–834, Jul. 2013.

[25] A. M. Goberman, "Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson disease," *Med. Sci. Monitor*, vol. 11, no. 3, pp. CR109–CR116, 2005.

[26] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. Vargas-Bonilla, K. Daqrouq, S. Skodda, J. Rusz, and E. Nöth, "Automatic detection of Parkinson's disease in running speech spoken in three different languages," *J. Acoust. Soc. Amer.*, vol. 139, no. 1, pp. 481–500, Jan. 2016.

[27] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease," *J. Acoust. Soc. Amer.*, vol. 129, no. 1, pp. 350–367, 2011.

[28] M. A. Little, P. E. McSharry, S. J. Roberts, D. A. Costello, and I. M. Moroz, "Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection," *Biomed. Eng. OnLine*, vol. 6, no. 1, p. 23, 2007.

[29] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1264–1271, May 2012.

[30] K. Kira and L. A. Rendell, "A practical approach to feature selection," in *Proc. 9th Int. Conf. Mach. Learn.*, 1992, pp. 249–256.

[31] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[32] N. Mlambo, W. K. Cheruiyot, and M. W. Kimwele, "A survey and comparative study of filter and wrapper feature selection techniques," *Int. J. Eng. Sci.*, vol. 5, no. 8, pp. 57–67, 2016.

[33] Y. Bengio, O. Delalleau, N. L. Roux, J. F. Paiement, P. Vincent, and M. Ouimet, "Spectral dimensionality reduction," in *Feature Extraction: Foundations and Applications*. Berlin, Germany: Springer, 2006, pp. 519–550.

[34] G. Roffo and S. Melzi, "Features selection via eigenvector centrality," in *Proc. New Frontiers Mining Complex Patterns (NFMCP)*, Oct. 2016, pp. 1–12.

[35] K. K. Bharti and P. K. Singh, "Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering," *Expert Syst. Appl.*, vol. 42, no. 6, pp. 3105–3114, Apr. 2015.

[36] P. Gomez, F. Diaz, A. Alvarez, K. Murphy, C. Lazaro, R. Martinez, and V. Rodellar, "Principal component analysis of spectral perturbation parameters for voice pathology detection," in *Proc. 18th IEEE Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2005, pp. 41–46.

[37] J. P. Teixeira, P. O. Fernandes, and N. Alves, "Vocal acoustic analysis–classification of dysphonic voices with artificial neural networks," *Procedia Comput. Sci.*, vol. 121, pp. 19–26, Jan. 2017.

[38] M. M. A. Maathidi, "Optimal feature selection and machine learning for high-level audio classification-a random forests approach," Ph.D. dissertation, Univ. Salford, Manchester, U.K., 2017.

[39] C.-F. Tsai and Y.-C. Chen, "The optimal combination of feature selection and data discretization: An empirical study," *Inf. Sci.*, vol. 505, pp. 282–293, Dec. 2019.

[40] Z. Ji and B. Wang, "Identifying potential clinical syndromes of hepatocellular carcinoma using PSO-based hierarchical feature selection algorithm," *BioMed Res. Int.*, vol. 2014, pp. 1–12, Mar. 2014.

[41] Z. Ji, G. Meng, D. Huang, X. Yue, and B. Wang, "NMFBFS: A NMF-based feature selection method in identifying pivotal clinical symptoms of hepatocellular carcinoma," *Comput. Math. Methods Med.*, vol. 2015, pp. 1–12, 2015.

[42] I. Razzak, R. A. Saris, M. Blumenstein, and G. Xu, "Integrating joint feature selection into subspace learning: A formulation of 2DPCA for outliers robust feature selection," *Neural Netw.*, vol. 121, pp. 441–451, Jan. 2020.

[43] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, "Benchmark for filter methods for feature selection in high-dimensional classification data," *Comput. Statist. Data Anal.*, vol. 143, Mar. 2020, Art. no. 106839.

[44] L. Ni, F. Fang, and J. Shao, "Feature screening for ultrahigh dimensional categorical data with covariates missing at random," *Comput. Statist. Data Anal.*, vol. 142, Feb. 2020, Art. no. 106824.

AMIRA S. ASHOUR received the B.S. degree in electrical engineering (electronics and electrical communication engineering) from the Faculty of Engineering, Tanta University, Egypt, in 1997, the M.Sc. degree in image Processing for nondestructive evaluation performance from the EEC Department, Faculty of Engineering, Egypt, in 2000, and the Ph.D. degree in smart antenna (direction of arrival estimation using local polynomial approximation) from the Faculty of Engineering, Egypt, in 2005. She has also been the Vice Chair of the Computer Engineering Department, Computers and Information Technology College (CIT), Taif University, Saudi Arabia for one year, since 2015. She was the Vice Chair of the Computer Science Department, CIT College, Taif University, Saudi Arabia for five years, until 2015. She has been an Assistant Professor and also the Head of the EEC Department, Faculty of Engineering, Tanta University, Egypt, since 2016. She has also been a member of the Research and Development Unit, Faculty of Engineering, Tanta University, Egypt, since August 2019. She has also been the Engineering Manager of Huawei ICT Academy, Tanta University, Egypt, since September 2019. She has 20 edited books and four authored books along with about 150 published journal and conferences articles. Her research interests include biomedical engineering, image processing and analysis, medical imaging, computer-aided diagnosis, signal/image/video processing, machine learning, smart antenna, direction of arrival estimation, targets tracking, optimization, and neutrosophic theory. She is an Editor-in-Chief for the *International Journal of Synthetic Emotions* (IJSE), IGI Global, US. She is an Associate Editor in several International Journals. She is Series Co-Editor of Advances in Ubiquitous Sensing Applications for Healthcare series, Elsevier.

MAJID KAMAL A. NOUR received the bachelor's degree in electrical Engineering (biomedical) from King Abdulaziz University, in 2007, the master's degree in biomedical engineering from La Trobe University, Australia, in 2010, and the Ph.D. degree in electronics engineering (biomedical) from the Royal Melbourne Institute of Technology (RMIT), Australia, in 2014. He is currently an Assistant Professor with King Abdulaziz University (KAU) and also an Active Researcher in nanotechnology, biomedical engineering and sensors with several highly cited publications. He is keen in Hospital design, medical equipment acquisition and commissioning, medical equipment regulation and standards. He is also a Medical Equipment and Hospital Design Consultant. He is also a member of the Saudi Scientific Society for Biomedical Engineering and the Saudi Society for Quality.

KEMAL POLAT received the B.Sc. and M.Sc. degrees from the Electrical-Electronics Engineering Department, Selcuk University, in 2002 and 2004, respectively, the Ph.D. degree in electrical and electronic engineering from Selcuk University, in 2008, and the master's degree from the Department of Electrical and Computer Engineering, University of Houston, in 2016. He designed mathematical modeling of memory performance in his postdoctoral work. He is currently a Professor Dr. with the Electrical and Electronic Engineering Department, Engineering of Faculty, Bolu Abant Izzet Baysal University, since July 2019. He has 112 articles published in the SCI journals and 65 international conference articles.). His google H-index is 36. He is in the 35th place in the world rank with 5367 citations according to the keyword Biomedical Signal Processing in Google Scholar. His research interests are biomedical signal classification, control systems, electronic, statistical signal processing, visual memory, neuroscience, brain–computer interface, PPG signal, medical electronics, digital signal processing, pattern recognition, and classification. He is a Member of the Editorial Board of the *Journal of Neural Computing and Applications* (SCI expanded) and *Applied Soft Computing*, Elsevier (SCI J). He has been an Associate Editor of *Measurement* journal at Elsevier, since 2019.

YANHUI GUO received the B.S. degree in automatic control from Zhengzhou University, China, the M.S. degree in pattern recognition and intelligence system from the Harbin Institute of Technology, China, and the Ph.D. degree from the Department of Computer Science, Utah State University, USA. He was a Research Follow with the Department of Radiology, University of Michigan, and also an Assistant Professor with St. Thomas University. He is currently an Assistant Professor with the Department of Computer Science, University of Illinois at Springfield. He has published more than 70 journal articles and 20 top conference articles, completed more than ten grant funded research projects. His research areas include computer vision, machine learning, big data analytics, computer aided detection/diagnosis, and computer assist surgery. He worked as an associate editor of different international journals and a reviewer for top journals and conferences.

WAFAA ALSAGGAF was an Assistant Professor with the Information Technology Department, College of Computer science and Information Technology King Abdulaziz University, in 2014, also a Supervisor of the Information Technology Department, College of Computer Science and Information Technology, Computer Skills Unit (Girls Section / Alsulaymaniyah), King Abdulaziz University, from 2016 to 2017, and also a Supervisor of the Information Technology Department, College of Computer Science and Information Technology, King Abdulaziz University, from 2017 to 2019. Her research areas include mobile and ubiquitous computing, mobile and ubiquitous learning, technology enhanced learning, human–computer interaction, and E-learning.

AMIRA EL-ATTAR received the B.S. and M.S. degrees in electronics and electrical communications engineering from the Faculty of engineering, Tanta University, in 2003 and 2013, respectively, and the Ph.D. degree in data management for Internet of Things in medical application from the Faculty of Engineering, Tanta University, Egypt, in 2019. From 2003 to 2013, she was a Teacher with the Department of Computer Science, Higher Institute of Management and Information Technology, Kafr EL Sheikh, Egypt. From 2013 to 2019, she became an Assistant Lecturer. Since September 2019, she has been a Lecturer with the Department of Computer Science, Higher Institute of Management and Information Technology.

· · ·