Received May 11, 2020, accepted May 29, 2020, date of publication June 10, 2020, date of current version June 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3001426

Multi-Objective Optimization of Wavelet-Packet-Based Features in Pathological Diagnosis of Alzheimer Using Spontaneous Speech Signals

MAHDA NASROLAHZADEH[1,2], JAVAD HADDADNIA[1], AND SHAHRYAR RAHNAMAYAN[2], (Senior Member, IEEE)

[1]Department of Biomedical Engineering, Hakim Sabzevari University, Sabzevar 9617976487, Iran

[2]Department of Electrical, Computer, and Software Engineering, Ontario Tech University, Oshawa, ON L1G 0C5, Canada

Corresponding author: Mahda Nasrolahzadeh (ms.nasrolahzadeh@yahoo.com)

ABSTRACT Alzheimer's disease (AD) ranks among the main types of neurodegenerative disorders. Patients suffering AD should tackle serious problems since their language skills malfunction. The impact of such disorders is reflected by reduced quality and feature variation of spontaneous speech signals in speech analysis. This paper aims at assessing the variations of some specific types of these energy- and entropy-based features within the frequency range of the speech signals. In the approach followed, the wavelet-packet coefficients are utilized to extract the energy and entropy measures at every spectral sub-band in six successive levels of decomposition. However, the decomposition process conducts a set of high-dimensional feature vectors that is a challenging task for feature selection. This study suggests the application of a Non-dominated Sorting Genetic Algorithm-II (NSGA-II) for enhancing a group of the sub-band indexes of a wavelet-packet for which the extracted features lead to the highest diagnosis rate of the grouping of Alzheimer's and healthy individuals. The technique proposed here showed that the best overall classification results for both optimized entropy feature vs. energy are more noticeable in discriminating patients with AD from healthy subjects. It is also confirmed the significant impact of multi-objective feature selection on performance of classification (i.e., disease diagnosis) and, its conformity to the disordered nature of the biological signals could help diagnose AD in an efficient manner.

INDEX TERMS Alzheimer's disease, spontaneous speech signal, wavelet packet, multi-objective optimization, feature selection, non-dominated sorting genetic algorithm-II.

I. INTRODUCTION

It is believed that patients will be detected quicker for more clinical trials to detect dementia earlier by developing non-invasive intelligent methods. Moreover, having improved systems with present objective analysis for automatic grouping and diagnosis of dementia could pave the way for visiting the patients in a timely manner for upcoming medical and financial choices. These techniques do not change or hamper the patients' abilities, since the spontaneous speech in these techniques is not supposed to be a stressful test by the patient. That is why non-invasive intelligent techniques that are cheap and more convenient for diagnosing dementia are significantly involved in clinical practices [1]–[3].

Alzheimer's disease is a progressive neurodegenerative disorder ranking among the most widespread kinds of dementia affecting older people [4], [5]. The disease is connected with reduced mental ability involving problems with memory, understanding, judgment, thinking, and language use. Besides memory loss, the loss of language skills is another significant issue brought about by AD. The loss of language-related communication ability by Alzheimer is dependent on the stage of the disease. These speech deficits can be separated into three stages: pre-clinical stage, intermediate stage, and advanced stage. The pre-clinical stage involves difficulties in finding the correct words in a spontaneous speech that is often not found. Furthermore, subject at this stage

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

cannot find the correct word to say what they mean. At the intermediate stage, daily language and vocabulary usage is limited with the answers sometimes very inadequate covering few words in the advanced stage [6]–[8]. The spontaneous speech is produced not only from a unique non-linear system with complex physiologic fluctuations, which has complex nonlinear interactions with some physiological subsystems but also it reflects the dynamical activity of the brain. Since AD patients have the biomechanical parameters of the disturbed airflow generated in the vocal tract undergo alteration. Actually, the altered biomechanical parameters accounts for the irregularities in their pattern of vibration [9], [10]. In fact, the main reason for the irregularities in their pattern of vibration is the change in their biomechanical parameters. These irregularities might include pitch frequency fluctuations, airflow volume changes, amplitude and mucosal wave reduction and the noise-like turbulence of airflow near the cords [11].

In fact, such changes in the waveform and the audible quality of patients' spontaneous speech signals are more obvious. As the resonance pattern of human speech as the convolution of the glottal pulse, excited by the airflow running through the vocal tract, may be an important indicator of the brain function. It would be possible to track the speech signal characteristic variations, which often assume the aspect of conditions that the nerve cells in the brain die or no longer function normally, to detect any neurodegenerative disorder.

Recent literature has paid significant attention to measure some fundamental parameters of disordered speech signals accurately. Hadjitodorov and Mitev presented a new parameter, normalized first harmonic energy (NFHE) for pathological estimation of vocal fold disorders [12]. Moreover, they introduced a computerized system for automatic analysis of pathological voices. K. López *et al.* combined the fractal dimension (FD) of the observed time series with linear parameters in the feature vector to diagnose AD earlier [13]. Hansen *et al.* used the differential Teager energy operator and the energy separation algorithms for estimating the vocal fold irregularities [14]. Nasrolahzadeh *et al.* [15] focused on the efficiency of non-linear dynamic measures like higher order spectra (HOS) and phase coupling for analyzing spontaneous speech signals spoken by normal subjects and AD patients. They argued that AD patients had decreased quadratic phase couplings of spontaneous speech signals compared to a healthy individual. Nevertheless, AD patients showed shifted speech phase coupled harmonics to the higher frequencies healthy individuals. Based on the bispectrum and bicoherence coefficients, they have also proposed some innovative techniques [16] for higher order spectra to extract discriminative information from the spontaneous speech signals of the healthy subjects and three groups of AD. In a previous study [17], to quantify and compare the contribution of nonlinear and chaotic dynamics of human speech variability in patients with AD and control subjects, Poincare plots of speech variability signals were analyzed. It was found that different stages of AD induced different reflections on speech features.

New studies on the techniques developed in the field of pathologic speech signal processing reported that they had a common goal. Because over repeated recordings, objective and subjective voice quality characteristics are different in normal and pathological speech signals, the objective is following the difference of a group of specific features in normal and patient subjects.

The literature shows several studies reporting wavelet-based processing as a robust tool to investigate non-stationary signals, which are useful for extracting speech properties [18]–[20]. The wavelet packet transform (WPT) provides information for the demonstration of a signal in the time-scale plane with a vast range of possibilities [21]. In this study, the theory of hierarchical wavelet-packet decomposition has more elaborately been utilized for investigating the variations of the signals of individuals with and without Alzheimer's. The loss of language skills in AD patient with Alzheimer's leads to varied expected level of energy within the frequency range of disordered speech signals. However, abnormality in the vibration pattern of human speech as the convolution of the glottal pulse, accounts for the variations in signals entropy. Nonetheless, because of poor signal during the spontaneous speech, the distribution manner of such alterations within the whole frequency range of pathological speech signals remains unclear. It only seems reasonable with no mathematical proof that these types of variations may follow a common rule in all the speech waveform of the patients diagnosed with a specific type of disorder. The spontaneous speech signals and spectrograms of a control subject and an Alzheimer's patient are shown in Fig. 1 in the respective order. Given the loss of language skills for speaking in AD, establishing the relationship with the natural environment, he suffers an important poverty (more and longer pauses or silence sections) in his signal during spontaneous speech. Consequently, difference of energy and entropy measures in frequency sub-bands of disordered signals is likely [22].

This study aimed to evaluate the ability of energy and entropy features to describe the origin of the loss of language skills, which is reflected in both difficulties in speaking and comprehension in AD. It has also been tried to investigate that either entropy or energy can provide a practical basis for AD diagnosis from the spontaneous speech signal. Nevertheless, the most effective feature can show the features of a signal in a group of less connected feature vectors using which the highest diagnosis performance would be obtained, so it can be used a firm basis to differentiate the classes automatically.

Furthermore, in order to design an elaborate diagnostic system from spontaneous speech signals, the ability of energy and entropy features is very important [23]. Consequently, these features are determined using several sub-bands of the speech signals of Alzheimer's and healthy subjects. The wavelet-packet coefficients at a different level of decomposition computed by the variation of energy and entropy features in frequency sub-bands of the speech signals will be used for characterizing the AD speech signal.

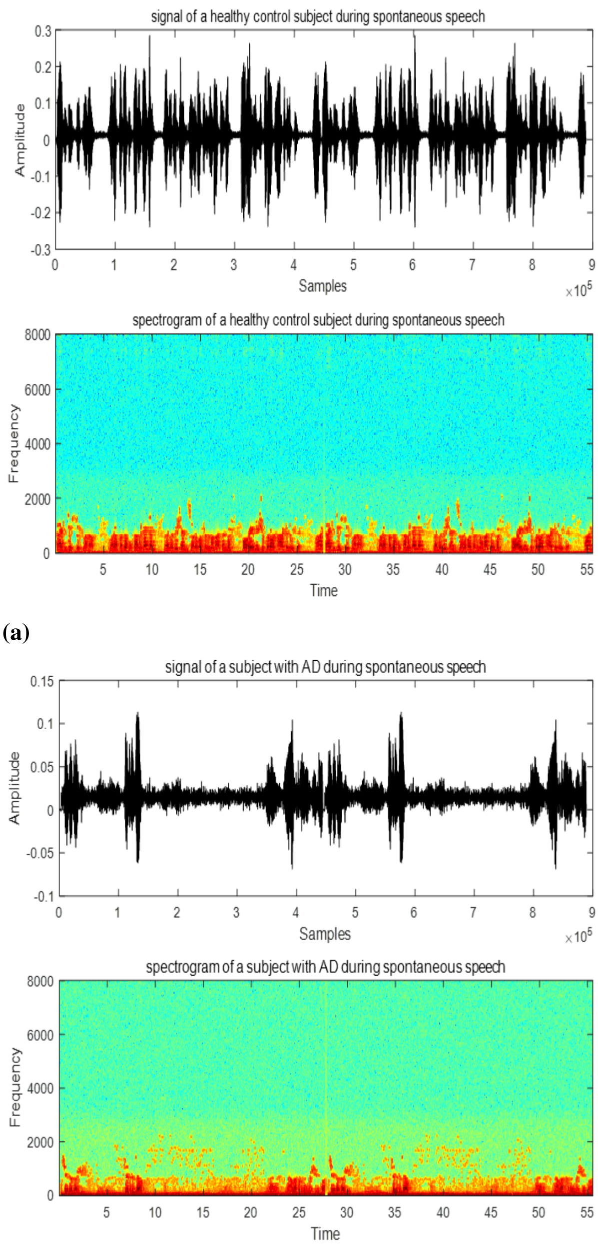


FIGURE 1. Representative signal and spectrogram of a healthy control subject (a) an Alzheimer's patient (TS) (b) during spontaneous speech. The signal and spectrogram are totally different in structure spontaneous speech signals. Because of loss of language skills for talking, understanding of the patient with Alzheimer's, the poverty of the speech signals for the patient with Alzheimer's is obvious [15].

It is worth mentioning that the flexibility achieved by the wavelet packet transform (WPT) decomposition has not yet been entirely investigated for feature extraction. If the non-redundant representations do not restrict the search for an optimal decomposition, the number of possible solutions to search will increase markedly leading to a difficult combinatorial problem.

For obtaining an optimal selection of a group of low-correlation frequency sub-bands that help the extracted features to distinguish individuals with and without Alzheimer's, a Non-dominated Sorting Genetic Algorithm-II (NSGA-II) is applied which has the ability to maximize the classification accuracy while minimizing the number of features that are extracted from frequency sub-bands [24]. In this study, to select suitable features for early diagnosis of AD and classification of patients with AD and control subjects, a classifier based on pattern recognition neural network (PRNN) [25] is utilized to evaluate the efficiency of selected solutions, using spontaneous speech signals of the Alzheimer's and control subjects. The proposed method, which is expressed as evolutionary wavelet packets (EWPs), uses the advantages offered by multi-objective evolutionary optimization to find a novel method for early diagnosis of AD based on energy and entropy extraction of features of Alzheimer's and control subjects from spontaneous speech signals. The schematic diagram of the proposed approach is shown in Fig. 2.

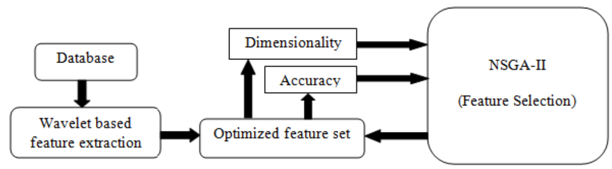


FIGURE 2. The schematic diagram of the proposed multi-objective optimization approach.

The rest of the paper is organized as follows: In section II, the dataset utilized in this study is briefly described, then, the methods and the quantification analyses used in the study are presented. In section III, the experiment results and performance of the proposed techniques are reported. In section IV, discussions are presented. Finally, section V presents the conclusions.

II. MATERIALS AND METHODS

A. DATABASE

A database was made according to the patients in the Old Nursing Home in Sabzevar, Razavi Khorasan Province, Iran. The selection of the participants was done based on the consensus diagnostic guidelines and according to availability of the special communication problems of the patients based on the relative stage of their disease for age-matched healthy normal individuals and AD patients [26]. The Institutional Review Board of all participants approved the study and it was conducted following the ethical standards of the 1964 Declaration of Helsinki. After explaining the aim of the study, all individuals provided us with their informed consent before any involvement.

30 individuals (52 – 98 years old, 14 women, 16 men) were diagnosed with Alzheimer's disorder, with three levels of AD: First Stage (FS), Second Stage (SS) and Third Stage (TS), (FS = 6, SS = 15, TS = 9) (AD group). The patients in the three stages of AD (namely, pre-clinical AD, mild cognitive

impairment (MCI) due to AD, and dementia due to AD) had a Mini-Mental State Examination (MMSE) score of 14–26 and a Clinical Dementia Rating (CDR) of 0.5 or 1.0 and fitted the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA) criteria for probable AD. 30 healthy control subjects (HCS), (52 – 98 years old, 15 women, 15 men), were selected as the control. The normal subjects were non-depressed, non-MCI, and non -demented and had an MMSE score of 27–30 and a CDR of 0.

The subjects were explicitly told to have a friendly conversation telling interesting personal stories and voice their feelings under a relaxed and friendly recording atmosphere was. An audio recorder was applied for recording the speech signals at the Old Nursing Home in Sabzevar. The audio was extracted in WAV format with sampling frequency and bit per second of 16 KHz and 16 Kbits, in the respective order. The speech signals were recorded about fifteen to seventeen hours for the control and AD groups, in the respective order. The recording time for AD subjected took less time since they needed more time for finding the words compared to the healthy individuals. They spoke more slowly, less evidently, with longer pauses. Their message was interrupted or unfinished with more time required for finding the words. Our previous works presented a more detailed explanation of the database [15], [27].

B. FEATURE EXTRACTION

The wavelet transform (WT) is used for extracting the features from the spontaneous speech signals. A given signal uses low-pass and high-pass filters, which H and L impulse response, to decompose the signal to detail (high frequency content) ($d_j(k)$) and approximation (low frequency content) ($a_j(k)$) coefficients. In other words, the signal is decomposed into a coarse detail and approximation information. The wavelet coefficients can be calculated using the following iterative relations:

$$d_j(k) = \sum_m H(m-2k)\, a_{j+1}(n) \tag{1}$$

$$a_j(k) = \sum_m L(m-2k)\, a_{j+1}(n) \tag{2}$$

The relation between these two filter banks of length N is determined using the following equation [21]:

$$H(n) = (-1)^n L(N-1-n) \tag{3}$$

As illustrated in Fig. 3, because of the orthogonal property, the wavelet-packet coefficients at various scales and positions allow representing the information included in a discrete signal by choosing various sub-trees from the full decomposition and is calculated by the following equations:

$$C^r_{n,k} = \sqrt{2}^r \sum_{m=-\infty}^{+\infty} f(m)\,.W_n(2^r m - k) \tag{4}$$

$$C^{r-1}_{2n,l} = \sum_m L(m-2l)\,.C^r_{n,m} \tag{5}$$

$$C^{r-1}_{2n+1,l} = \sum_m H(m-2l)\,.C^r_{n,m} \tag{6}$$

In which r is the scale index, l is the translation index, L and H are the impulse response of the low-pass and high-pass filters connected with the approximation and detail coefficients, in the respective order. The relation between these two filter banks of length K can be given by [28]:

$$H(m) = (-1)^m L(K-1-m) \tag{7}$$

The wavelet packet is can show some variation of features in signal without any redundancy or overlap of information between the signals of different sub-bands. We used this significant feature for describing the properties of a signal. The wavelet-packet decomposition procedure of the spontaneous speech signals of a control subject and an Alzheimer's patient is shown in Fig. 3.

For a group of wavelet-packet coefficients of a signal, the measures of energy in each of the frequency bands have been computed as:

$$Energy_n = \frac{1}{n^2} \sum_{k=1}^{n} |C^r_{n,k}|^2 \tag{8}$$

As discussed above, the orthogonality involves no redundancy or overlap of information between the signals of different sub-bands. This property enables us to search the whole frequency range of the speech signal for a group of optimal feature vectors, which can better differentiate Alzheimer's and healthy group speech signals.

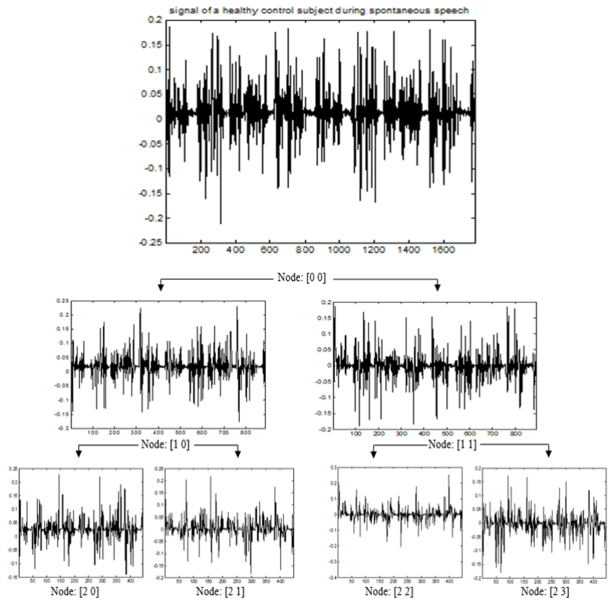
Recently, some researchers used entropy feature on the speech processing for different purposes [29], [30]. It was reported that entropy is significantly involved in the information theory as a descriptor of information, choice, and uncertainty [31]. In this study, the irregular pattern variations of the pathological effect of Alzheimer's speech were characterized by entropy. In this study, the irregular pattern variations of the pathological effect of Alzheimer's speech were characterized by entropy. The Shannon entropy indicates that how the cross-correlation value of pathological speech signals in the time domain is affected by the produced irregularities in the variations of AD speech [32]. Furthermore, it was reported as the top orthogonal basis that can be determined using the extracted wavelet-packet coefficients as follows:

$$Entropy_n = -\sum_{k=1}^{n} |C^r_{n,k}|^2 log |C^r_{n,k}|^2 \tag{9}$$

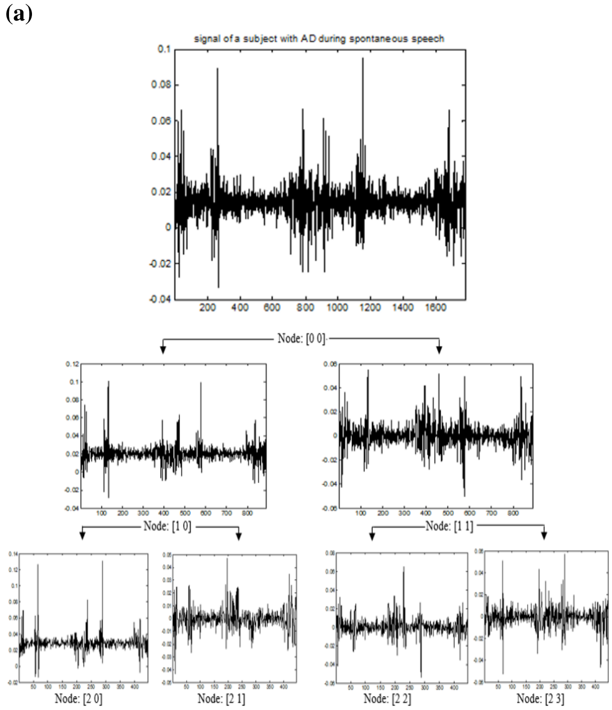
C. CLASSIFICATION PROCEDURE

Studies show that pattern recognition neural network (PRNN) is a good classifier in the pathological diseases, such as [33], [34] and other biological signals for predicting and forecasting disease in multiple areas [35]–[37]. This study used PRNN for classifying the extracted wavelet-packet-based energy and entropy features of AD, and HC groups from spontaneous speech signals.

In the PRNN, the selection of sufficient neurons in the hidden layer of a two-layer feed-forward network allows the



(a)



(b)

FIGURE 3. Two level wavelet-packet decomposition of spontaneous speech signals, (a) a healthy control subject and (b) an Alzheimer's patient. The calculated wavelet packets decomposition at different scales and positions gives an index of the performance of the given measures. Because of the orthonormal property, such a tree is yielded for a discrete signal.

vectors to be properly classified with the sigmoid hidden and Softmax output neurons.

In this paper, PRNN was trained using the Levenberg–Marquardt (LM) back-propagation algorithm [38], [39]. The LM algorithm is designed based on the second-order learning speed approach without calculating the Hessain matrix (H). When the function is a sum of squares, then the Hessian matrix can be approximated based on Jacobian matrix [40]. The LM algorithm uses the following equation to approximate the Hessian matrix:

$$x_{k+1} = x_k - [J^{\tau} J + \mu I]^{-1} J^{\tau} e \tag{10}$$

When μ is considered zero, this algorithm only uses the Newton-Gauss method to calculate the Hessian approximation, and when μ is a very large number, this relation becomes a gradient reduction method with a small step.

This study investigated some structures of PRNN, hidden layer and a number of its neurons. Furthermore, all the input patterns were separated into 3 parts based randomization: 70% of input patterns are assigned to the training set, 15% of them are allocated to the validation set and the rest to the testing set [15]. After yielding the best topology and convenient number of training epochs using the validation set, the blind testing patterns are applied to the classifier and assessing the performance was done. Furthermore, the performance of the proposed method was further explored using specificity, sensitivity, and accuracy metrics [41]. Next, the training set was used to train the pattern recognition network followed by testing our proposed method using the test set.

D. NSGA-II WITH MULTIPLE OBJECTIVES

The meta-heuristic algorithms inspired by the natural process of evolution, have been widely applied for finding the global optima in complex search spaces [42]–[44]. These algorithms require an objective function to perform the search according to the problem under study. Actually, in real-world optimization, problems are infrequently confronted with one objective and are often encountered with two or more conflicting objective functions subject to several certain constraints. The optimal solution of multi-objective optimization (MOPs) is not a single solution but a set of solutions defined as the Pareto-optimal front [45]. Furthermore, the best solution requires the trade-off between objectives. For addressing the multi-objective problems by using evolutionary algorithms and obtaining the Pareto optimal set and, therefore, non-dominant set of solutions, each objective is considered as a limitation or the combinations of the individual objective functions into a single collective function [46]. In other words, a set of candidate solutions is Pareto optimal, for which none of the objectives can be enhanced with no deterioration of at least another objective function. A MOPs problem can be defined as:

$$\begin{aligned} &\text{Max/Min} : F(\vec{x}) = f_1(\vec{x}), f_2(\vec{x}), \ldots, f_o(\vec{x}) \\ &\text{Subject to} : \begin{cases} g_i(\vec{x}) \geq 0, & i = 1, 2, \ldots, m \\ h_i(\vec{x}) = 0, & i = 1, 2, \ldots, p \\ L_i \leq x_i \leq U_i, & i = 1, 2, \ldots, n \end{cases} \end{aligned} \tag{11}$$

where o ($o \geq 2$) is the number of objectives, F indicates the vector of objectives to be enhanced, g indicates the set of feasible solutions linked with inequality constraints, h indicates the set of feasible solutions related with equality

TABLE 1. Parameter's setting used in NSGA-II algorithm [24].

parameters	value
Maximum Number of Iterations	30
Population size	100
Crossover Percentage	0.8
Mutation Percentage	0.4
Mutation Rate	0.1
Number of Runs	10

constraints, and L and U are the lower and upper bound for each decision variable x_i, in the respective order.

For finding the Pareto front in the (MOPs) problems, many modifications and changes have been used to the classical genetic algorithm (GA). In [47], a variation of the classical multi-objective algorithm was proposed, namely, the fast non-dominated sorting genetic algorithm (NSGA-II) that is a popular fast elitist multi-objective, non-domination version of the GA and is an instance of an evolutionary algorithm from the evolutionary computation algorithm. NSGA-II can find well-spread solutions over the Pareto-optimal front while necessitating a low computational complexity as follows:

$$O\left(mN^2\right) \tag{12}$$

where N and m are the population size and the number of objectives, in the respective order. The chief components of NSGA-II consist of a fast non-dominated sorting approach, a fast-crowded distance estimation procedure, and ultimately, a simple crowded comparison operator. Algorithm 1 and Table 1 show the NSGA-II algorithm and parameter settings used in the proposed method, in the respective order.

Algorithm 1 NSGA-II Algorithm

Generate uniform random initial population W_j
Calculate objective values for each individual in W_j using Eq. (13)
Allot rank based on Pareto dominance for each individual in W_j
Generate offspring population O_j
for $j = 1$ to number of generation **do**
  for all Parent and offspring population **do**
    Allot rank and crowding distance based on Pareto
    Create sets of non-dominated fronts
    Specify the crowding distance among the stronger individual
  end for
  if $j < generation$ **then**
    $j = j + 1$
    $W_j = W_{j-1} \cup O_{j-1}$
  else
    Stop the loop and product O_j
  end if
end for

E. FINDING OPTIMAL WAVELET FEATURES USING NSGA-II

In the suggested NSGA-II, the first objective function E assesses the selected feature subset leading to a measure of classification performance. A PRNN classifier is applied by the NSGA-II for assessing the solutions during the search [48]. This means that the PRNN-based spontaneous speech classifier is utilized as the first objective function and the classification accuracy is achieved for each evaluated individual. The selection operator signifies a group of individuals offering the maximum fitness values. Then, these groups are chosen for bearing the children of the next generation. In the PRNN-based detection rate process, both the mutation and crossover operators are accountable for creating next-generation members. The PRNN classifier is trained on a corpus of human speech signals for two groups: AD individuals and healthy subjects, and the accuracy obtained on a test set is the return value of the first objective function. Removing the irrelevant and redundant features is desired for obtaining a solution with higher performance. Consequently, the second objective function n_f considers the number of selected features, preferring smaller subsets. The trade-off between these two objective functions in the feature optimization process to reach a set of solutions as Pareto optimal solutions are defined as:

$$Z = E \times \left(1 + \beta n_f\right) \tag{13}$$

A set of solutions is achieved, as Pareto optimal solutions offering a trade-off following the NSGA-II algorithm run. In such solutions, any improvement in one objective function degrades the other objective function. In other words, choosing a solution with a small number of features decreases the computational cost and restricts the performance of the classification resulting in an increased error rate. For instance, Figs. 5 and 7 show the Pareto fronts yielded using E and n_f as objective functions for both energy and entropy features, respectively. Based on both functions, the best individuals in the population move in a direction to the ideal optimum. In this approach, all steps covered in the WP feature selection scheme and Algorithms 2 and 3, in the respective order, shows the information for the population evaluation.

Algorithm 2 Optimization For WP-Based Feature

Obtain WPT for a corpus of spontaneous speech signals for two groups; HC and AD groups using Eqs. (5), (6), (8) and (9)
Initialize the NSGA-II population
Calculate W_j (**Algorithm 3**)
repeat
  Choose parents
  Generate a new population W_j (j = 1) from the parents chosen
  Substitute population
  Calculate W_j
  until a stopping criteria is reached

Algorithm 3 Calculate Fitness Based on Classification

for each individual in W_j **do**
Re- initialization train/test vectors according to the chromosome
Train the PRNN based classifier on the train set
Test the PRNN based classifier on the test set
Appoint classification rate as the current individual's fitness
end for

III. EXPERIMENTS AND RESULTS

This paper aimed at evaluating the impact of optimization techniques on the applicability of our proposed method. To achieve this aim, an experimental study was conducted on the 120 segments database described in section II-A, which includes 30 AD patients and 30 healthy controls. As discussed above, all the speech signals were pre-processed. For performing the wavelet-packet decomposition of the speech signals to their time-frequency localized coefficients, segmenting the spontaneous speech signals was required. Our previous works offered a detailed description for calculating the segmentation and the processing on the speech signals [27], [49]. Selecting an efficient wavelet function and the number of decomposition levels in the analysis of signals is of tremendous importance. Consequently, wavelet families were compared to know which one is the most suitable for spontaneous speech signals of healthy control subjects and AD patients [50]. According to literature, the wavelet families including Haar, Daubechies, Meyer, Symmlets, and Coiflets y Splines were studied [51], [52]. As mentioned above, speech signals were decomposed into six levels. Then, the low-frequency component of the signals was reconstructed. Next, the mean and maximum of errors for both the original and reconstructed signals were calculated. Finally, the best wavelet base was selected in accordance with the errors. Fig. 4 depicts the maximum error and mean error of the reconstructed spontaneous speech signal. As a result, the first order Daubechies or Haar WT for the optimization experiments was chosen. The features are the energy and Shannon entropy of the wavelet-packet coefficients, at the six levels of decomposition.

In this research, spontaneous speech signals of patients with AD, and HC subjects are decomposed in order to create the lower and higher frequency bands. In such a decomposition procedure, for dividing the signal spectrum into a specific number of sub-bands, the wavelet-packet low-pass, and high-pass filter-banks are used repetitively. Consequently, a 64-dimensional feature space, at the sixth decomposition level, would have been searched for the solution of the optimization problem.

NSGA-II was applied on the collected data by choosing the feature vector optimization for evaluating the ability of features in distinguishing the two groups; the classification performance was also assessed. In this process, the NSGA-II searches for increased detection rate in a group of least correlated feature vector indicators. The population size of the NSGA-II was set to 100. The NSGA-II offers a set of individuals dominating the actual population at the end of every generation. This means, other individuals are not closer to the Pareto front. Then, the chromosome is chosen that yielded the best accuracy from the optimal set offered in the last generation.

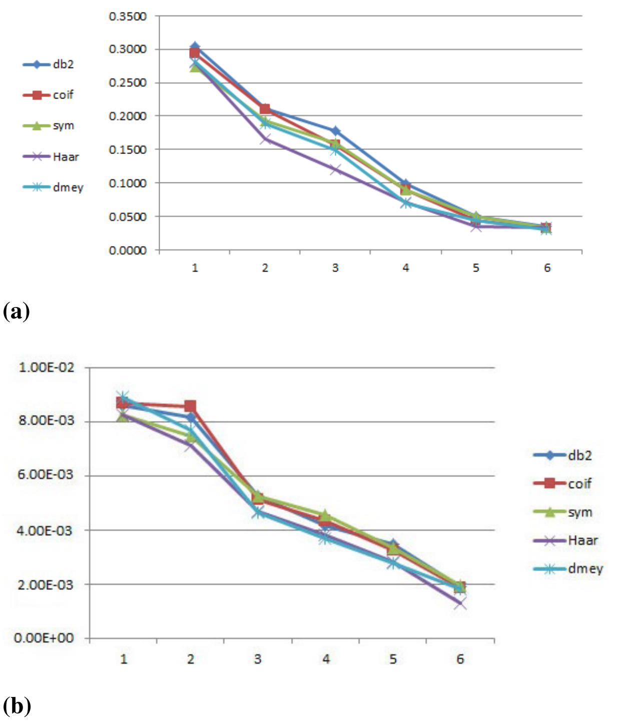


FIGURE 4. The maximum error and mean error of the reconstructed spontaneous speech signal. (a) Maximum error (b) Mean error.

We run the suggested technique according to Algorithm 2 and two energy and Shannon entropy Pareto fronts were obtained. The algorithm is run 30 times independently. For each case study based-feature optimization process, the combination of 30 Pareto fronts into one union set is also performed. The classification of subjects into the AD and HC groups is performed by PRNN. To test the optimization of the experiments, the efficiency of the classification of optimized feature sets was evaluated. To this end, the selected optimal features were fed into the classifier for classification. Afterwards, the feature vector optimization process, in most of the decomposition levels enhanced the recognition rates in each case study.

A. CASE STUDY-I: ENERGY FEATURE VECTORS OPTIMIZATION PROCESS

For validating the optimization results, the first optimization problem was done for showing the impact of energy feature vectors of the decomposition levels allowing the maximization of the classification accuracy while minimizing the number of features. The result of the distribution of Pareto

TABLE 2. The results Pareto fronts of applying NSGA-II to the energy feature vectors.

Energy	Pareto fronts (1).Out	Pareto fronts (3).Out
S	[1, 4, 5]	[6]
n_f	3	1
r_f	0.4286	0.1429
E	0.3466	0.3466

S is the groups of the selected feature, n_f is the number of selected features, r_f is the percentage of the selected feature, and E is the amount of cross-entropy error between the features.

optimal solutions of applying NSGA-II to the energy feature vectors for spontaneous speech signals of two groups, AD, and healthy subjects is shown in Fig. 5. These solutions show the highest possible classification accuracy and the lowest number of features that can be reached through the current design (see Fig. 5). As shown in Table 2, the result of utilizing NSGA-II to the energy feature indicates that the quality of the solutions obtained through this algorithm is remarkable. Consequently, after running of 30 iterations, two sub-bands were selected for energy measure. Furthermore, it could produce consistent results over many runs irrespective of the initial randomized population. Additionally, choosing a compromising solution admissible by the choice of the wavelet-packet decomposition is possible since the yield solution set via the proposed algorithms involves a spectrum of optimal solutions [50].

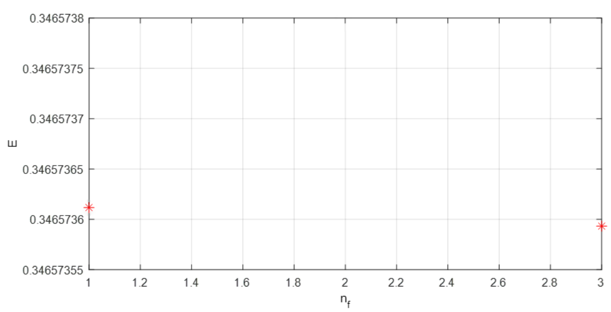
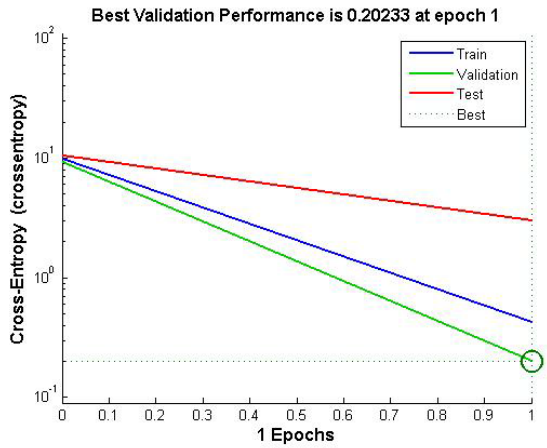


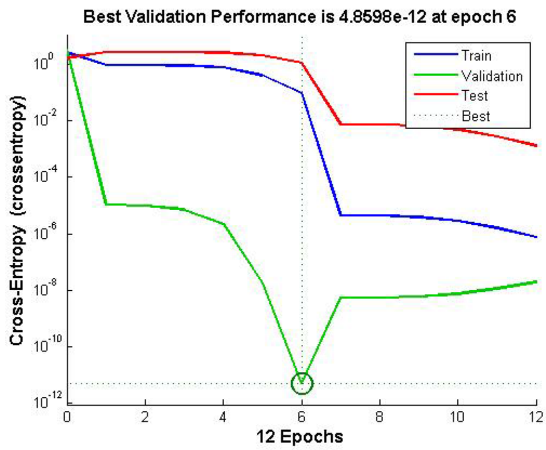
FIGURE 5. Pareto fronts of energy feature obtained from NSGA-II experiment.

In the optimization process, NSGA-II yields a set of individuals dominating the real population at the end of each generation showing that no other individual is close to the Pareto front in the process. Then, the chromosome was selected which obtained the best accuracy from the optimal set provided in the previous generation. The classification abilities of the optimized feature set were also calculated for each of the optimization the experiments. This evaluation was conducted using a ten-fold cross-validation procedure, through the evaluation data set, 230 composed of all the energy feature vectors of spontaneous speech signals for individuals with and without Alzheimer's. For this purpose, the data is first to divided into ten parts or folds. Next, validation and training are repeated 10 fold; one fold of the data is kept out in each repetition for validating the classifier, while the first nine are applied for training [53].

The classification results for groups of AD and HC were achieved when the energy feature was applied with three and one sub-bands of wavelet-packet decomposition, in the respective order. The classification results are shown in Table 3 and Fig. 6.



(a)



(b)

FIGURE 6. The performance of data training and testing of PRRN for energy-based features. (a) with [1, 4, 5] of sub-bands (b) with [6] of sub-band.

B. CASE STUDY-II: SHANNON ENTROPY FEATURE VECTORS OPTIMIZATION PROCESS

For exploring the impact of Shannon entropy feature vectors of spontaneous speech signals of the Alzheimer's and healthy

TABLE 3. Classification results AD vs. HC with energy feature.

No. of sub-bands	Sensitivity (%)	Specificity (%)	Accuracy (%)
[1, 4, 5]	95	100	97.5
[6]	91.7	100	95.8

subjects in the optimal classification accuracy and number of features is carried out using NSGA-II. The results of utilizing NSGA-II to the Shannon entropy features are shown in Fig. 7, and Table 4, respectively. As shown, the process of optimizing the number of selected feature vectors has resulted in the enhancement of the detection rates. According to the results of Pareto fronts analysis of the Shannon entropy feature, the quality of the solutions achieved through the proposed method is similar to those in the energy feature.

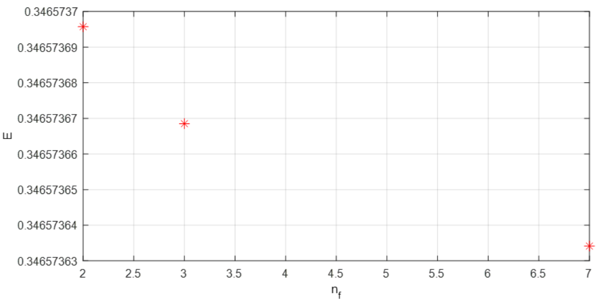
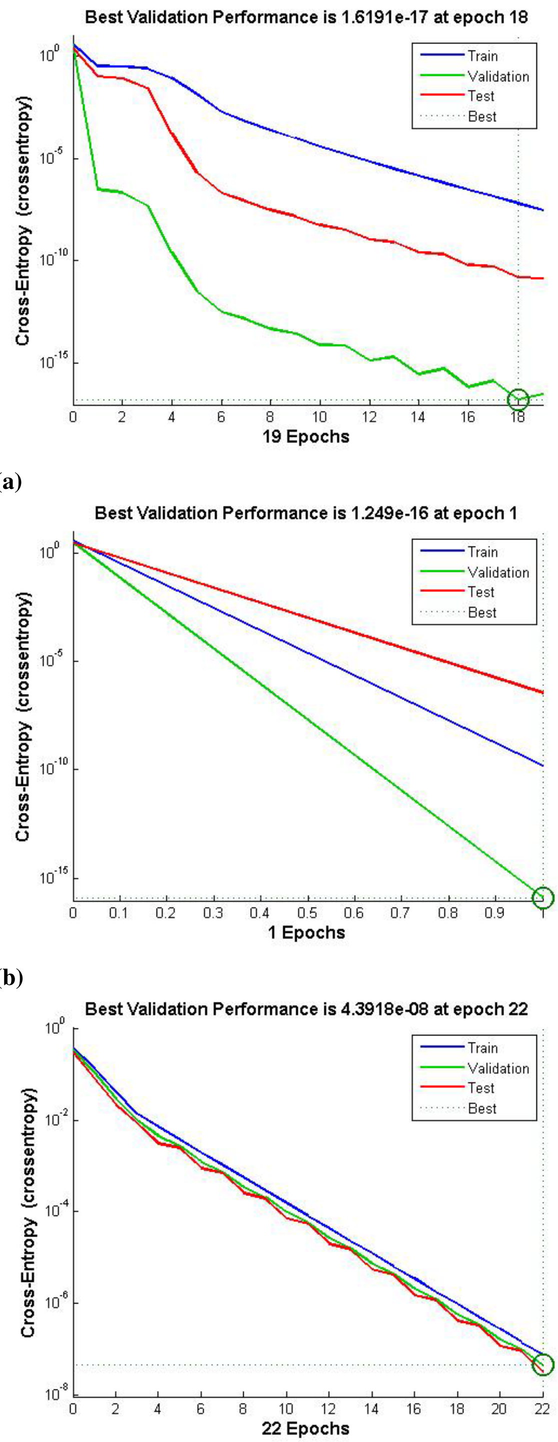


FIGURE 7. Pareto fronts of Shannon entropy feature obtained from NSGA-II experiment.

The classification results for groups of AD and HC were obtained when Shannon entropy was used with seven, two, and three sub-bands of wavelet-packet decomposition, respectively. The results are shown in Table 5, Fig. 8, respectively.

According to Tables 2 and 4, the results of two Pareto fronts are significantly different. Consequently, after running of 30 iterations, two sub-bands were selected for energy measurement while three and two sub-bands, respectively, were selected for Shannon entropy measurement. As shown, the analysis of the Shannon entropy and energy Pareto fronts produces better results. Nonetheless, the Shannon entropy feature outperformed the energy feature in the AD diagnosis. It was also found to be clinically significant. As previously mentioned, the feature vector space is searched by NSGA-II to find the best individual. In this process, a set of individuals dominating other individuals of population found by NSGA-II show that no other individual is closer to the Pareto front. Next, from the optimal set achieved in the previous generation, the chromosome was selected for the best accuracy. Furthermore, the increased level of wavelet decomposition led to the extraction of a more complex group of feature vectors resulting in increased detection rate. This shows that both features are beneficial in distinguishing healthy signals from AD signals and they can affect differentially the features of spontaneous speech signals during Alzheimer's stages. It is



(c)

FIGURE 8. The performance of data training and testing of PRRN for Shannon entropy-based features. (a) with [1, 2, 3, 4, 5, 6, 7] of sub-bands (b) with [3, 6] of sub-bands (c) with [3, 5, 7] of sub-bands.

worth mentioning that the accurate assessment of energy and Shannon entropy features may enable us to find some of the new aspects of the signals, which can be expressly used to explain the AD and healthy signals. Besides, according to

TABLE 4. The results Pareto fronts of applying NSGA-II to the Shannon entropy feature vectors.

Shannon entropy	Pareto fronts (1).Out	Pareto fronts (2).Out	Pareto fronts (3).Out
S	[1, 2, 3, 4, 5, 6, 7]	[3, 6]	[3, 5, 7]
n_f	7	2	3
r_f	1	0.2857	0.4286
E	0.3466	0.3466	0.3466

S is the groups of the selected feature, n_f is the number of selected features, r_f is the percentage of the selected feature, and E is the amount of cross-entropy error between the features.

TABLE 5. Classification results AD vs. HC with Shannon entropy.

No. of sub-bands	Sensitivity (%)	Specificity (%)	Accuracy (%)
[1, 2, 3, 4, 5, 6, 7]	89.85	94.12	91.67
[3, 6]	98.41	94.73	96.67
[3, 5, 7]	96.77	100	98.33

Tables 2 and 4, the Shannon entropy feature outperformed the energy feature in AD diagnosis efficiently.

According to Figs. 6 and 8, the results of accuracy, plus the rate of convergence of the PRNNs model for the Shannon entropy feature are superior to the energy feature in terms of the classification of the two mentioned groups. It is obvious that the best results were obtained when the sub-band of the Shannon entropy based features was selected as [3, 5, 7]. Ultimately, according to these plots, the features are diverse, and can be used for discriminating the HC and AD groups with greater accuracy. However, according to the results of simulations in this study, the maximum detection rates have been achieved at all level of wavelet packet decomposition for both energy and entropy measures. Furthermore, the PRNNs detection rates have been higher for entropy-based features vectors in all the wavelet packet levels of decomposition.

This paper explored the AD effects for assessing the performance reliability of both features of energy and entropy of spontaneous speech signals in the specific psychophysiological state. As discussed above, language skills are influenced by AD and these variations are shown by the patients' spontaneous speech signals. Therefore, it expectedly changed some measures such as the energy and the entropy of the signals. Although such disturbances certainly alter the characteristics of the speech signal, the distribution function of these changes is not clear. In other words, there is not a model, which is able to describe the variation of the characteristic in different spectral frequency signals. Actually, the impact of loss of language skills leading to in difficulty in both speaking and comprehension during Alzheimer's stages might be significant in a group of feature vector indicators of the extracted from different spectral frequencies. However, the optimal feature subsets extracted for Alzheimer's and control subjects have the lowest level of cross-correlation. Consequently, the spectral distribution of energy and entropy measures are distinguished for individuals with and without AD. Since entropy can assess the degree of irregularity in the signal waveform, the most important reason for the success of this feature is.

Another significant fact in the present study is that the feature selection was defined as a bi-objective optimization problem minimizing both the number of features and the error rate of classification. By applying the NSGA-II, the method of optimal feature sub-set selection yielded to a group of spectral sub-bands, which are scattered all over the frequency range of the speech signals. Moreover, in NSGA-II, the evaluation and, as a result, chromosomes guidance is carried out based on their readiness in the objectives space. In fact, an evolutionary process occurred in the objectives space. In this process, the problem space could be improved the evolutionary process of algorithm in terms of objectives (the classification accuracy and the number of features), in feature selection problem, consequently, it was essential to be extended well to involve the vast areas of the Pareto front uniformly.

Table 6 provides a comparison between the results of two case study based feature optimization processes. As shown, the analysis of the Pareto Front in both feature dataset shows a remarkable performance by NSGA-II.

Furthermore, obviously, the Shannon entropy-based features outperform the energy-based features regarding the classification of the two mentioned groups. The best accuracy with the energy based features was achieved by the sub-bands [1, 4, 5], and the best accuracy with the Shannon entropy based features was achieved with the sub-bands [3, 5, 7], in the respective order. Furthermore, AD stage was detected by the PRNN classifier with the Shannon entropy based features with an accuracy index of 98.33 %. Therefore, the energy and entropy features extracted from the spontaneous speech signals are appropriate descriptors of abnormalities in the speech of the healthy individuals, and AD patients. Thus, a more effective scheme to diagnose AD early is proposed by the introduced approach.

Given the performance of the overall system, the optimization of the decomposition based on the WPT led to highly accurate diagnosis and this is a significant point to diagnose AD early. The effectiveness of NSGA-II as a multi-objective evolutionary optimization algorithm, in the early diagnosis of AD and classification of AD patients and HC subjects using spontaneous speech signals is confirmed by our results. However, a precise comparison is not possible because the majority of studies used different databases and study populations. Furthermore, different studies applied different classifiers and common types of cross-validation approaches, e.g. leave-one-out or K-fold cross-validation. Therefore, for this purpose, three more popular classifiers were used to

TABLE 6. Classification results energy vs. Shannon entropy in the feature optimization process.

Feature	No. of sub-bands	Sensitivity (%)	Specificity (%)	Accuracy (%)
Energy	[1, 4, 5]	95	100	97.5
	[6]	91.7	100	59.8
Shannon entropy	[1, 2, 3, 4, 5, 6, 7]	89.85	94.12	91.67
	[3, 6]	98.41	94.73	96.67
	[3, 5, 7]	96.77	100	98.33

further investigate the proposed method by applying the same feature extraction method from the viewpoint of classification accuracy and diagnostic, and to compare them with each other for assessing the performance of the pipeline. The Decision Tree (DT), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) were utilized. A detailed description of these classifiers can be found in [54]–[56]. The performance of the classifiers was evaluated using ten-fold cross-validation [53]. Table 7 shows the results of the classification for the three different classifiers and our proposed method.

As can be seen, the best accuracy (98.33%) was obtained by the PRNN classifier. Besides, the DT classifier yielded an accuracy index of 97.2%, which is remarkable. The worst accuracy (96.51%) yielded by the SVM with the Shannon entropy features. However, the KNN was able to detect Alzheimer's with a classification accuracy of 96.7%. Moreover, the best specificities and sensitivities are relevant to the PRNN while the worst specificities and sensitivities are relevant to the SVM. Thus, the PRNN classifier yields better performance than three other classifiers in this study.

IV. DISCUSSION

Human speech is produced when the air is pushed out from the lungs up the vocal tract. The vocal cords produce a periodic signal by filtering out the white noise signals produced by the lungs. The periodic signal produced so is further filtered by other parts of the human vocal apparatus. The ultimate sound produced in this way is called voiced speech. The frequency of human speech is determined by the resonance pattern of the sound. However, unvoiced speech is white noise like the sounds produced by the lips and teeth. In human speech, unvoiced speech and voiced speech are combined to produce phonemes and words [57]. This process results in a regular pattern of energy discharge. The result of such arrangements is recorded as a clear periodic signal during the opening and closing stages of the glottal.

As explained before, the effect of speech disability of people with Alzheimer's disease can be more explicitly investigated as a group of disorders caused by abnormal brain changes. Through the speech signals of patients, the negative impact of such pathologies is also reflected. Because of the loss of language skills reflected in talking, understanding, and relationship with the natural environment, there can be significantly observed in the signal for a person with Alzheimer's during spontaneous speech. Therefore, the variations of some quantities could be envisaged such as the entropy and the energy of the signals. Although the distribution function of these variations is not clearly established, such irregularities certainly alter the characteristics of the speech signal. In other words, it is not feasible to figure out a model to demonstrate the characteristic changes in the signals of various spectral sub-bands. However, it is worth mentioning that the effect of the loss of language skills in Alzheimer's disease, while being examined in more detail in various spectral sub-bands, may be significant in a set of extracted feature vector indices. In addition, while elicited for Alzheimer's and healthy subjects, the lowest levels of cross-correlation contain for such optimal feature sub-sets. Therefore, with respect to such optimal feature sub-sets, the distinction between energy and entropy measures in terms of the spectral distribution in the healthy subjects and those with Alzheimer's would be understandable.

The main reason for the successful result of the Shannon entropy in the early diagnosis of Alzheimer's disease is that it can evaluate the measure of disorder in the signals' waveform. Such disorders in AD speech can originate from brain damages that cause the typical symptoms of frontotemporal dementia containing variations in cognitive abilities and problems with language use, there can be clearly observed important poverty in his signal during spontaneous speech [27], [58].

Another notable finding of the present study is that the optimal feature subset selection method yields a set of spectral sub-bands that are scattered all over the frequency ranges of the speech signals, which is performed using NSGA-II. In other words, the frequency range of such determinants features exceeds the range of speech signals in several conventional implementations. The frequency range of the features utilized is not limited to the low and medium frequency range of spontaneous speech signals in the early diagnosis of Alzheimer's disease. Note that the influence of the loss of language skills reflected on change of speech signal properties not only affects the low and medium frequency sub-bands of the Alzheimer's speech but also it severely induces variability of such features in the high-frequency range of spontaneous speech signals of subjects with AD.

As far as we know, this is the first study designed with the aim of early diagnosing AD by examining the piecewise variation of some specific types of features, known as energy and Shannon entropy using structural spontaneous speech data. As discussed above, AD influences the language skills leading to the reduced quality and feature variation in speaking. The innovation of this approach is highlighting the role of Shannon entropy features, in comparison with

TABLE 7. Classification results of the proposed method using four classifiers.

classifiers	Specificity (%)	Sensitivity (%)	Accuracy (%)
DT	97.48	96	97.2
KNN	97.31	95.22	96.7
SVM	96.82	94.05	96.51
PRNN (proposed method)	100	96.77	98.33

TABLE 8. The proposed method in comparison with other studies.

Reference	year	Feature extraction technique (number of features)	Classification technique	Database	Accuracy (%)
K.Lopez et.al [[59]]	2015	Fractal dimension and duration features (100)	MLP classifier	private	97.7
A.Konig [[60]]	2015	First vocal markers (6)	SVM classifier	private	81
Nasrolahzadeh et al. [[49]]	2015	Acoustic features (100)	Fuzzy classifier	private	97.96
Nasrolahzadeh et al. [[15]]	2018	Higher Order Spectral based features (5)	DT classifier	private	97.71
Mirzaei et al. [[61]]	2018	Temporal and acoustical voice features (39)	KNN classifier	private	94
Our proposed method	2020	Energy- and entropy- based features (2)	PRNN classifier	private	98.33

energy, as a basis for shedding light on the speech differences between the healthy subjects, and those with Alzheimer's. Overall, our findings show that Shannon entropy feature can evaluate and measure such differences in spontaneous speech. Furthermore, this feature is easy to calculate, and highly accurate in discriminating Alzheimer's and healthy subjects. Table 8 summarizes the comparison between our proposed method and the other provided systems in terms of diagnosis and classification results. As shown, the proposed method outperformed other studies regarding diagnostic performance.

V. CONCLUSION

This study introduces a new method to optimize the wavelet-based feature in spontaneous speech signals for early diagnosis of AD in AD patients and control subjects, using evolutionary computation methods for probing large and complex search spaces. In fact, it was tried to develop a multi-objective strategy that can maximize the discrimination capability and reliability of the overall system performance while minimizing the number of features. The energy and entropy features were extracted for six successive levels of wavelet-packet decomposition. NSGA-II was used for the optimal features selection at every level of decomposition reaching an excellent trade-off between redundancy and dimensionality for performing powerfulness in spontaneous speech classification. The experimental results indicated that the space of the optimized features increased the separation of the classes. Consequently, the proposed method, as an optimum design methodology for selecting robust speech features, was capable of boosting the classification performance showing that it could provide a good choice for the early diagnosis of AD. In future work, it would be interesting to analyze this feature selection method with more datasets are required to further demonstrate the reliability and potential power of our proposed method.

REFERENCES

[1] F. Previtali, P. Bertolazzi, G. Felici, and E. Weitschek, "A novel method and software for automatically classifying Alzheimer's disease patients by magnetic resonance imaging analysis," *Comput. Methods Programs Biomed.*, vol. 143, pp. 89–95, May 2017.

[2] J. Lee, Y. Kim, Y. Jeong, D. L. Na, J.-W. Kim, K. H. Lee, and D. Lee, "Inference of brain pathway activities for alzheimer's disease classification," *BMC Med. Inform. Decis. Making*, vol. 15, no. S1, p. S1, 2015.

[3] G. Fiscon, E. Weitschek, G. Felici, P. Bertolazzi, S. De Salvo, P. Bramanti, and M. C. De Cola, "Alzheimer's disease patients classification through eeg signals processing," in *Proc. IEEE Symp. Comput. Intell. Data Mining (CIDM)*, Dec. 2014, pp. 105–112.

[4] H. Braak and E. Braak, "Neuropathological stageing of Alzheimer related changes," *Acta Neuropathologica*, vol. 82, no. 4, pp. 239–259, 1991.

[5] A. E. Budson and P. R. Solomon, *Memory Loss E-Book: A Practical Guide for Clinicians*. Amsterdam, The Netherlands: Elsevier, 2011.

[6] J. Dauwels, F. Vialatte, and A. Cichocki, "Diagnosis of Alzheimer's disease from EEG signals: Where are we standing?" *Current Alzheimer Res.*, vol. 7, no. 6, pp. 487–505, 2010.

[7] F. Martínez-Sánchez, J. J. G. Meilán, E. Pérez, J. Carro, and J. M. Arana, "Patrones de prosodia expresiva en pacientes con enfermedad de Alzheimer," *Psicothema*, vol. 24, no. 1, pp. 16–21, 2012.

[8] W. Hu, C. McMillan, D. Libon, S. Leight, M. Forman, V.-Y. Lee, J. Trojanowski, and M. Grossman, "Multimodal predictors for Alzheimer disease in nonfluent primary progressive aphasia," *Neurology*, vol. 75, no. 7, pp. 595–602, 2010.

[9] P. Gómez, J. I. Godino, F. Rodríguez, F. Díaz, V. Nieto, A. Álvarez, and V. Rodellar, "Evidence of vocal cord pathology from the mucosal wave cepstral contents," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, May 2004, p. V-437.

[10] P. Gómez, J. I. Godino, F. Díaz, A. Álvarez, R. Martínez, and V. Rodellar, "Biomechanical parameter fingerprint in the mucosal wave power spectral density," *currents*, vol. 1, no. M1, p. f1, 2004.

[11] K. Ishizaka and N. Isshiki, "Computer simulation of pathological vocal-cord vibration," *J. Acoust. Soc. Amer.*, vol. 60, no. 5, pp. 1193–1198, 1976.

[12] S. Hadjitodorov and P. Mitev, "A computer system for acoustic analysis of pathological voices and laryngeal diseases screening," *Med. Eng. Phys.*, vol. 24, no. 6, pp. 419–429, 2002.

[13] K. López-de Ipina, J. Solé-Casals, H. Eguiraun, J. B. Alonso, C. M. Travieso, A. Ezeiza, N. Barroso, M. Ecay-Torres, P. Martinez-Lage, and B. Beitia, "Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach," *Comput. Speech Lang.*, vol. 30, no. 1, pp. 43–60, 2015.

[14] J. H. L. Hansen, L. Gavidia-Ceballos, and J. F. Kaiser, "A nonlinear operator-based speech feature analysis method with application to vocal fold pathology assessment," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 3, pp. 300–313, Mar. 1998.

[15] M. Nasrolahzadeh, Z. Mohammadpoory, and J. Haddadnia, "Higher-order spectral analysis of spontaneous speech signals in Alzheimer's disease," *Cognit. Neurodyn.*, vol. 12, no. 6, pp. 583–596, 2018.

[16] M. Nasrolahzadeh, Z. Mohammadpoory, and J. Haddadnia, "A novel method for early diagnosis of Alzheimer's disease based on higher-order spectral estimation of spontaneous speech signals," *Cognit. Neurodyn.*, vol. 10, no. 6, pp. 495–503, 2016.

[17] M. Nasrolahzadeh and J. Haddadnia, "Poincaré plots of spontaneous speech signals during alzheimer's disease," *Mitteilungen-Saechsischer Entomologen*, vol. 119, pp. 358–365, Apr. 2016.

[18] M. T. Sadiq, X. Yu, Z. Yuan, Z. Fan, A. U. Rehman, G. Li, and G. Xiao, "Motor imagery eeg signals classification based on mode amplitude and frequency components using empirical wavelet transform," *IEEE Access*, vol. 7, pp. 127678–127692, 2019.

[19] C. Zhuang and P. Liao, "An improved empirical wavelet transform for noisy and non-stationary signal processing," *IEEE Access*, vol. 8, pp. 24484–24494, 2020.

[20] G. Fiscon, E. Weitschek, A. Cialini, G. Felici, P. Bertolazzi, S. De Salvo, A. Bramanti, P. Bramanti, and M. C. De Cola, "Combining eeg signal processing with supervised methods for Alzheimer's patients classification," *BMC Med. Inform. Decis. Making*, vol. 18, no. 1, p. 35, 2018.

[21] I. Cohen, S. Raz, and D. Malah, "Adaptive time-frequency distributions via the shift-invariant wavelet packet decomposition," in *Proc. IEEE-SP Int. Symp. Time-Frequency Time-Scale Anal.*, Oct. 1998, pp. 645–648.

[22] V. Mehdizadehfar, F. Almasganj, and F. Torabinezhad, "Investigation of the effects of speech signal length on vocal disorder sorting done via dynamic pattern modeling," *J. Voice*, vol. 31, no. 4, p. 515-e1, 2017.

[23] E. S. Fonseca, R. C. Guido, S. B. Junior, H. Dezani, R. R. Gati, and D. C. M. Pereira, "Acoustic investigation of speech pathologies based on the discriminative paraconsistent machine (DPM)," *Biomed. Signal Process. Control*, vol. 55, Jan. 2020, Art. no. 101615.

[24] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[25] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

[26] I. G. McKeith, "Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): Report of the consortium on DLB international workshop," *Neurology*, vol. 47, no. 5, pp. 1113–1124, 1996.

[27] M. Nasrolahzadeh, Z. Mohammadpoori, and J. Haddadnia, "Analysis of mean square error surface and its corresponding contour plots of spontaneous speech signals in Alzheimer's disease with adaptive Wiener filter," *Comput. Hum. Behav.*, vol. 61, pp. 364–371, Aug. 2016.

[28] B. C. Sidney, *Introduction to Wavelets and Wavelet Transforms: A Primer*. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[29] H. Misra, S. Ikbal, H. Bourlard, and H. Hermansky, "Spectral entropy based feature for robust ASR," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 1, May 2004, p. I-193.

[30] H. Misra, S. Ikbal, S. Sivadas, and H. Bourlard, "Multi-resolution spectral entropy feature for robust ASR," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 1, Mar. 2005, p. I-253.

[31] S. Aviyente, "Information processing on the time-frequency plane," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 2, May 2004, p. II-617.

[32] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[33] Q. Yang, X. Zou, R. You, Y. P. Liu, Y. Han, Y. N. Zhang, L. Guo, H. Q. Mai, C. M. Xie, L. Li, and M. H. Hong, "Proposal for a new risk classification system for nasopharyngeal carcinoma patients with post-radiation nasopharyngeal necrosis," *Oral Oncol.*, vol. 67, pp. 83–88, Apr. 2017.

[34] J. M. Flores-Fernández, E. J. Herrera-López, F. Sánchez-Llamas, A. Rojas-Calvillo, P. A. Cabrera-Galeana, G. Leal-Pacheco, M. G. González-Palomar, R. Femat, and M. Martínez-Velázquez, "Development of an optimized multi-biomarker panel for the detection of lung cancer based on principal component analysis and artificial neural network modeling," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10851–10856, 2012.

[35] T. Tuncer, S. Dogan, and F. Ertam, "Automatic voice based disease detection method using one dimensional local binary pattern feature extraction network," *Appl. Acoust.*, vol. 155, pp. 500–506, Dec. 2019.

[36] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in pap smears," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 539–556, 2014.

[37] S. M. Vieira, L. F. Mendonca, G. J. Farinha, and J. M. C. Sousa, "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients," *Appl. Soft Comput.*, vol. 13, pp. 3494–3504, Aug. 2013.

[38] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. De Jesús, *Neural Network Design*, vol. 20. Boston, MA, USA: PWS, 1996.

[39] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *J. Soc. Ind. Appl. Math.*, vol. 11, no. 2, pp. 431–441, 1963.

[40] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, Nov. 1994.

[41] P. Langley, W. Iba, and K. Thompson, "An analysis of Bayesian classifiers," in *Proc. 10th Nat. Conf. Artif. Intell.*, vol. 90, Jul. 1992, pp. 223–228.

[42] C. D. Lin, C. M. Anderson-Cook, M. S. Hamada, L. M. Moore, and R. R. Sitter, "Using genetic algorithms to design experiments: A review," *Qual. Rel. Eng. Int.*, vol. 31, no. 2, pp. 155–167, 2015.

[43] G. Fiscon, E. Weitschek, E. Cella, A. L. Presti, M. Giovanetti, M. Babakir-Mina, M. Ciotti, M. Ciccozzi, A. Pierangeli, P. Bertolazzi, and G. Felici, "MISSEL: A method to identify a large number of small species-specific genomic subsequences and its application to viruses classification," *BioData Mining*, vol. 9, no. 1, 2016, Art. no. 38.

[44] A. A. F. Neto, A. M. Canuto, and J. C. Xavier-Junior, "Hybrid metaheuristics to the automatic selection of features and members of classifier ensembles," *Information*, vol. 9, no. 11, p. 268, 2018.

[45] K. Deb, "Multi-objective optimization," in *Search Methodologies*. Boston, MA, USA: Springer, 2014, pp. 403–449.

[46] C. A. C. Coello, "Multi-objective evolutionary algorithms in real-world applications: Some recent results and current challenges," in *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences*. Cham, Switzerland: Springer, 2015, pp. 3–18.

[47] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multiobjective optimization: Formulationdiscussion and generalization," in *Proc. ICGA*, vol. 93. New York, NY, USA: Citeseer, Jul. 1993, pp. 416–423.

[48] I. Salman, O. N. Ucan, O. Bayat, and K. Shaker, "Impact of metaheuristic iteration on artificial neural network structure in medical data," *Processes*, vol. 6, no. 5, p. 57, May 2018.

[49] M. Nasrolahzadeh, Z. Mohammadpoori, and J. Haddadnia, "Optimal way to find the frame length of the speech signal for diagnosis of Alzheimer's disease with PSO," *Asian J. Math. Comput. Res.*, vol. 2, no. 1, pp. 33–41, Feb. 2015.

[50] Y. Long, L. Gang, and G. Jun, "Selection of the best wavelet base for speech signal," in *Proc. Int. Symp. Intell. Multimedia, Video Speech Process.*, Oct. 2004, pp. 218–221.

[51] S. Mallat, *A Wavelet Tour of Signal Processing*. Amsterdam, The Netherlands: Elsevier, 1999.

[52] H. L. Rufiner and J. Goddard C, "A method of wavelet selection in phoneme recognition," in *Proc. 40th Midwest Symp. Circuits Syst.*, vol. 2, Aug. 1997, pp. 889–891.

[53] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of Database Systems*. 2009, pp. 532–538.

[54] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.

[55] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[56] N. Cristianini and J. Shawe-Taylor, *Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[57] B. Gold, N. Morgan, and D. Ellis, *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*. Hoboken, NJ, USA: Wiley, 2011.

[58] J. D. Rohrer, "Structural brain imaging in frontotemporal dementia," *Biochimica et Biophysica Acta (BBA)-Mol. Basis Disease*, vol. 1822, no. 3, pp. 325–332, 2012.

[59] K. Lopez-de Ipiña, J. Alonso-Hernández, J. Solé-Casals, C. M. Travieso-González, A. Ezeiza, M. Faundez-Zanuy, P. M. Calvo, and B. Beitia, "Feature selection for automatic analysis of emotional response based on nonlinear speech modeling suitable for diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 150, pp. 392–401, Feb. 2015.

[60] A. König, A. Satt, A. Sorin, R. Hoory, O. Toledo-Ronen, A. Derreumaux, V. Manera, F. Verhey, P. Aalten, P. H. Robert, and R. David, "Automatic speech analysis for the assessment of patients with predementia and alzheimer's disease," *Alzheimer's Dementia, Diagnosis, Assessment Disease Monitor.*, vol. 1, no. 1, pp. 112–124, 2015.

[61] S. Mirzaei, M. El Yacoubi, S. Garcia-Salicetti, J. Boudy, C. Kahindo, V. Cristancho-Lacroix, H. Kerhervé, and A.-S. Rigaud, "Two-stage feature selection of voice parameters for early alzheimer's disease prediction," *IRBM*, vol. 39, no. 6, pp. 430–435, 2018.

MAHDA NASROLAHZADEH received the M.Sc. degree (Hons.) in electronics engineering from Hakim Sabzevari University, Sabzevar, Iran, in 2015, where she is currently pursuing the Ph.D. degree with the Department of Biomedical Engineering. Since 2019, she has been a Visiting Research Scholar with Ontario Tech University, Oshawa, Canada. Her research interests include computer-aided diagnosis system design for the diagnosis of Alzheimer's disease with a focus on spontaneous speech signal and mathematical modeling, and optimization techniques.

JAVAD HADDADNIA received the B.Sc. degree in electrical and electronic engineering and the M.Sc. and Ph.D. degrees in electrical engineering from the Amirkabir University of Technology, Tehran, Iran, in 1993, 1996, and 2002, respectively. He has been working with the Department of Electrical and Computer Engineering, Hakim Sabzevari University, Sabzevar, Iran, since 2003. He is currently a University Professor. He has authored the book *The New Trends in the Application of Thermography Science for Diagnostic Purpose* (Supreme Century, USA, 2016) and has authored more than 150 articles in the artificial intelligence and digital signal processing. His current research interests include digital signal processing, artificial intelligence, machine vision, pattern recognition, and medical engineering. He received the Full University Professorship, in 2017, the Hakim Sabzevari University Award for Excellence in Scholarship Research, and the Creative Activity, from 2003 to 2005 for three years and 2009 to 2015 for five years. He received the Best Researcher Award from the Ministry of Science, Research and Technology, Iran, in 2006. He was a recipient of the Best Designer Award for National SCADA system for electrical power distribution network from the Ministry of Energy, Iran, in 2015.

SHAHRYAR RAHNAMAYAN (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in software engineering, and the Ph.D. degree in the field of evolutionary computation and optimization from the University of Waterloo (UW), Canada, in 2007. His Ph.D. thesis was proposed in the Opposition based Differential Evolution (ODE) and received more than 1060 citations, since 2008. Inspired from ODE, more than 500 articles have been published so far. Before joining to the Faculty of Engineering and Applied Science, Ontario Tech University, Canada, as a Tenure-Track Faculty Member, in 2008; he was a Postdoctoral Fellow with Simon Fraser University (SFU), Canada. He conducted research as a Visiting Associate Professor for two years at Michigan State University, MI, USA, from 2014 to 2016. He was granted tenure earlier and also was promoted to an Associate Professor position in 2013. His research interests include large-scale and multi-objective optimization, evolutionary computation, image processing, machine learning, and opposition-based soft computing. He is an active reviewer for more than 30 international conference and journal papers. He has published 160 articles mostly in optimization area, which received more than 5400 citations (H-index: 30). He was awarded the Ontario Graduate Scholarship (OGS), the President's Graduate Scholarship (PGS), the NSERC's Japan Society for the Promotion of Science (JSPS) Fellowship, the NSERC's Industrial Research and Development Fellowship (IRDF), the NSERC's Visiting Fellowship in Canadian Government Laboratories (VF), and the Canadian Institute of Health Research (CIHR) Fellowship for two times. He has been awarded several prestigious research grants, including the NSERC Discovery Grant and the Applied Research and Commercialization Initiative Fund. He has received the Ontario Tech University Research Excellence Award, in 2017.