International Journal on Advanced Science Engineering Information Technology

Vol.11 (2021) No. 1
ISSN: 2088-5334

Development of Intelligent Parkinson Disease Detection System Based on Machine Learning Techniques Using Speech Signal

Mohammed Younis Thanoun[a,1], Mohammad Tariq Yaseen[a,2], A.M.Aleesa[b,3]

[a]Department of Electrical Engineering, University of Mosul, Mosul, Nineveh, 41002, Iraq
[b]Department of Electrical & Electronic Engineering, UTHM, Parit Raja, Johor Baru, 86400, Malaysia
Corresponding author: [1]myounisth@uomosul.edu.iq, [2]mtyaseen@uomosul.edu.iq, [3]ahm3d.aleesa@gmail.com

Abstract— **Parkinson's disease is a brain condition that induces difficulty walking, standing, concentrating, trembling, and weakness. Parkinson's symptoms typically begin slowly and increase with time. Whenever the condition develops, individuals can experience trouble walking and communicating to others. Old people mostly tend to suffer from this disease and the number is expected to increase in the future. Machine learning (ML) techniques could help in the medical field in processing and analyzing data that offer good solutions in this field in terms of high accuracy and less required time compared to conventional methods. In this study, we proposed an enhanced methodology based on utilizing SMOTE to balance the dataset, due to the available dataset is imbalanced. then adopted extra tree classifier with k-fold technique after we balanced the dataset with SMOTE. we have achieved the best accuracy with respect to the classification accuracy in the literature, the obtained accuracy of our proposed model was higher than the used approaches in the related works. The new model for classifying the Parkinson's disease-dataset with class-imbalance data distribution achieved an accuracy of 96.52% by using our proposed method. The result shown that the dataset is lacked of balancing and it proves that the balancing in the dataset is important specially in medical classification. The impact of Optimal function selection, either automated by PCA or manually carried out, is clearly still being studied, and plays an essential role in improving the performance of machine learning.**

Keywords— **Parkinson's disease; machine learning; PD, SMOTE; extra tree classifier.**

Manuscript received 11 Jun. 2020; revised 7 Nov. 2020; accepted 10 Dec. 2020. Date of publication 28 Feb. 2021.

I. INTRODUCTION

Parkinson's Disease (PD) is recognized as a progressive neurodegenerative disease that has no cure yet. Different regions in the human brain, whether related to dopamine or not, are damaged by this disease. Older people mostly tend to suffer from this disease, and the number is expected to increase in the future. Even though many studies have been done to understand more about this disease, there is no cure for it as the particular reason that causes the disease is still unidentified. At the same time, statistics show that the number of infected individuals with PD is increasing. Motor and non-motor symptoms are signs of PD. Vital clinical signs of PD are impaired posture instability, resting tremor, bradykinesia, and rigidity [1]-[3].

As studies showed, men are expected to be more affected by PD than women. So, the rate of males got infected with PD is higher than females. The probability of being affected by PD is due to many reasons such as ethnicity, heredity, and polluted environment. One of the common known PD evaluation methods is Unified Parkinson's Disease Rating Scale (UPDRS) [4]. This method can be used to identify the development and risk of PD symptoms. It is based on a score resulting from the neurological assessment, which the specialist does. It is a subjective score that lacks sensitivity. Early detection of PD is very beneficial for many reasons, for instance, clinical and prognostic issues. As a result, many researchers have collected real data to predict this disease [5], [6].

Machine learning (ML) techniques could help the medical field process and analyze data that offer good solutions in this field in terms of high accuracy and less required time than conventional methods [7]. Utilizing the ML techniques needs an important parameter, that is, the dataset. Dataset is required data that is collected about the problem being studied. Different ML approaches and different datasets have been used in the medical field [8]. Accuracy of disease diagnosis and required time for the detection process are always the main challenging parameters and researchers trying to improve [9].

Regarding PD, Many ML techniques have been suggested and used for PD detection. Accuracy and processing time are still the main concern due to the big data for PD detection [10]. For instance, for PD progression, Least Square Support Vector Machine (LS-SVM) showed good performance as a regression method [11]. Features reduction for a hybrid system, classification, and clustering methods were introduced. It indicated that the incorporation of feature reduction or selection, feature pre-processing methods , and classification provides a classification accuracy of 100% [12]. The random Forest algorithm was tested and compared with other methods. Results showed that it is suitable for PD detection [13]. A fuzzy based nonlinear transformation method was proposed. Support Vector Machine (SVM) and Principal Component Analysis (PCA) were used to extract the optimal subset of features [14]. Meta-cognitive radial basis function network by Projection-based learning method was proposed for PD prediction [15]. The expectation-maximization algorithm and Genetic programming were combined to developed a hybrid system for PD detection [16]. Parallel neural networks have been suggested for PD detection. The performance results showed more than 8 % enhancement on the system performance than a single neural network [17]. Clustering machine learning and Incremental support vector machine methods were applied as an intelligent system for PD detection. The Unified Parkinson's Disease Rating Scale (UPDRS) detection accuracy, for Total and Motor , was 0.4656 and 0.4967, respectively [18].

For efficient PD detection, several types of classification methods were evaluated. It was shown that the overall classification score for the neural network is improved to be 92.9% [19]. Several methods such as General Regression Neural Networks, Support Vector Machine, Least Square Support Vector Machine, and Multilayer Perceptron Neural Networks were tested. The decision trees method was used for PD detection from voice data features. The obtained accuracy was 90% [20]. A fuzzy k-Nearest Neighbor method was applied for PD detection. Principal Component Analysis was applied to increase the accuracy to be 96.07% [6]. For Genetic Algorithm-Wavelet Kernel-Extreme Learning Machine (GA-WK-ELM), the accuracy of using this method for PD detection was 96.81% [21]. Support Vector Machine classifier was applied to get overall accuracy of 97% for the PD detection [22]. Particle swarm optimization (PSO) enhanced fuzzy k-Nearest Neighbor technique was introduced as a PD diagnosis system. The obtained accuracy of this technique was 97.47% [23]. Expectation-maximization (EM) algorithm and principal component analysis (PCA) were employed with Adaptive Neuro-Fuzzy Inference System (ANFIS) and Support Vector Regression (SVR) models for PD detection. The obtained accuracy of this technique was 98.20% [8]. Feature reduction, clustering, and classification algorithms were proposed as a hybrid system for PD detection. The achieved accuracy of this proposed system was 100% [12].

However, PD causes speech impairments for patients. Speech impairments are one of the earliest and significant symptoms for the early diagnosis of PD. Therefore, speech and voice signals investigation and analysis have drawn a lot of attention for PD prediction. Different studies are related to the classification techniques and methods of PD voice data sets. Several of these studies are listed here. The speech and voice symptoms frequency of 200 PD patients were recorded, categorized, and examined. Decision trees were used to classify the voice samples of healthy and patient individuals. The classified voice samples were examined using a threshold-based method to determine PD patients [20], [24]. A hybrid Artificial Intelligence based classifier method was proposed for early identification of PD. Lagrangian Support Vector Machine (LSVM) and Multi-Layer Perceptron (MLP) were used first for features selection then classification. The overall accuracy was 100% [25]. Voice impairment (dysphonia) was tested using Support Vector Machine (SVM), Random Tree (RT), and Feedforward Back-propagation based Artificial Neural Network (FBANN) ML classifiers. FBANN classifier has achieved the best performance with 97.37% recognition accuracy [26]. Automatic Relevance Determination (ARD) was incorporated with Gaussian process for effective feature selection. The accuracy based on voice samples of 42 PD patients was 96.92% [27].

Synthetic Minority Over-Sampling Technique (SMOTE) method was employed as a data preprocessing step to convert imbalanced dataset to balanced class distribution dataset [28]. After that, Random Forests classification technique was employed for PD classification of 756 attributes (192 healthy and 564 patient). The obtained accuracy by using only Random Forests classification was 87.037%. Meanwhile, the achieved accuracy of using a hybrid method of Random Forests classification and SMOTE was 94.89%.

High rate of vocal disability tends to happen due to early stages of PD. This has attracted researchers' attention to putting more effort into studying and analyzing speech signals to detect PD early. Tunable Q-factor wavelet transform (TQWT) was introduced for feature extraction from voice signals [29]. This study presented the efficiency of using TQWT compared with the up-to-date feature extraction methods used in the detection of PD. The voice signals in this study were collected from 252 individuals (188 patient and 64 healthy). The extracted feature subsets were applied to multiple classifiers in combination with ensemble learning approaches. TQWT showed comparable results to the recent voice signal processing methods employed for PD classification. Combining Mel-frequency cepstral (MFC) and TQW coefficients via filter feature selection technique provided complementary information of PD classification, improving the overall system performance. The obtained accuracy in this study was 86%.

The contribution was to improve over the obtained results in previous study [29]. In addition to improvements in terms of the used classifier and accuracy [28], the 754 features, as the whole feature sets, extracted from 252 patients in the dataset for PD detection based on voice signals. This work's proposed methodology, which is based on utilizing SMOTE due to the dataset with an imbalanced dataset then adopted an extra tree classifier with k-fold technique after we balanced the dataset with SMOTE.

II. Materials and Method

A. Dataset

The sample of the dataset for Parkinson's disease is from the UCI network [30]. This dataset includes 753 features and 756 samples. This dataset is also a problem of two classes: 564 Parkinson patients and 192 healthy trials. This dataset is thus unbalanced. The 2018 Parkinson disease dataset was developed by Sakar et al [29]. Attributes from voice signal processing techniques have been collected. Data from 188 PD patients (107 males and 81 females) between the ages of 33 to 87 have been taken from the PD dataset [29]. Figure 1 describes the Parkinson disorder sample type distribution. Two groups, one green (majority) and one red (minority). This graph reveals two sections. There is a great deal of inequality amongst these groups.

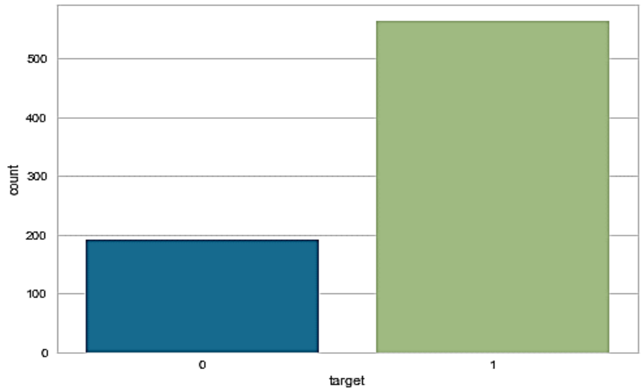

Fig. 1 dataset classes distribution

B. The Proposed Approach

This research is a new approach to identify the PD dataset using the SMOTE and Extra Trees Classifier in an unbalanced class distribution. First step, the problem of unbalanced dataset was handled using the SMOTE method in the PD dataset. PD datasets have then been transformed into a representation of the balance class with SMOTE. The block diagram of this hybrid system is seen in Figure 2. Second step is to use the balanced dataset with Extra Trees Classifier in order to detect the PD patients in the dataset.

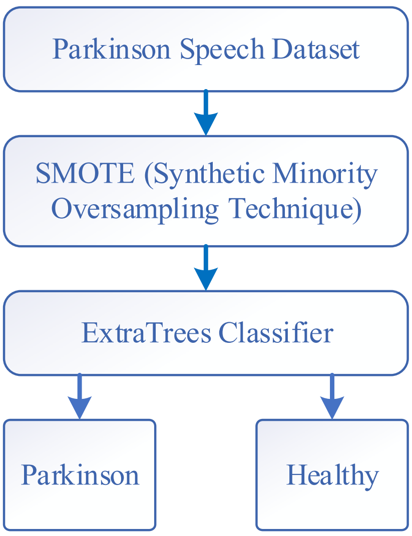

Fig. 2 Flowchart architecture for Parkinson classification

C. SMOTE

Synthetic Minority Over-sampling Technique (SMOTE) is a modern technique for over-sampling the minority community through processing, then de-sampling, artificial specimens [31]. One of the latest approaches is SMOTE, which is used to resolve class inequality issues in machine learning. Figure 3 displays the SMOTE algorithm schematically.

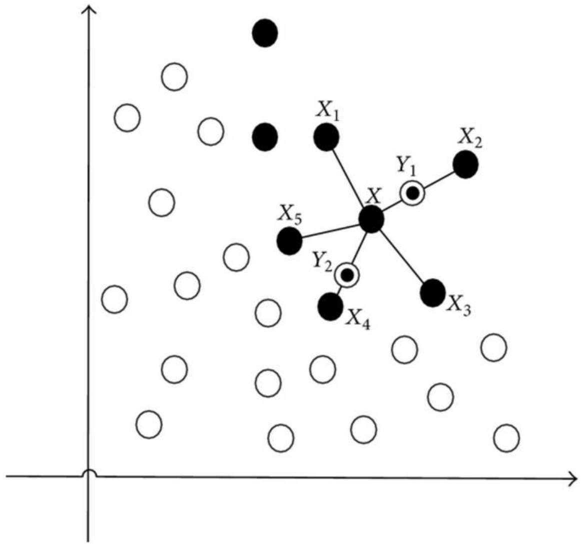

Fig. 3 the SMOTE algorithm representation diagram [32]

The new class distribution of the PD dataset as seen in Figure 4 applied the SMOTE Approach to PD datasets. As seen in the below figure, using the SMOTE synthetic experiments process, the number of small group (healthy) samples has increased. The problem of the unbalanced class was dealt with in the PD dataset classification using SMOTE algorithms. In total, 192 tests in the PD data collection have a healthy group and Parkinson's disorder category have 564 tests. The number of stable population samples in the PD dataset has increased and improved from 192 to 564 samples since adding the SMOTE approach to the PD dataset. There have been no changes in the number of samples in the Parkinson disease group.

```
After OverSampling, counts of label '1': 564
After OverSampling, counts of label '0': 564
```

Fig. 4 classes after applied SMOTE

D. Extremely Randomized Trees Classifier (ExtraTrees Classifier)

Extremely Randomized Trees that is known as ExtraTrees classifier which used in Compared to random decision forests, this distinction is distinct from how randomness is applied through training. Multiple trees are learned to build an extra Trees-classifier, each tree is trained in all training information. The optimal split in a node is identified in a subset of all possible features through analysis, and is equivalent to the random decision tree. A specific threshold for each feature is randomly selected instead of looking for each feature's right threshold. The one that corresponds to the highest improvement in the score is chosen from such random groups. During preparation, the greater degree of randomness generates more individual trees and reduces the variance further [33]. Because of these additional leaves, there are far stronger performance than random forests. 5% of training data was collected at random in order to reduce the training time for the classifier training. The classifier is then built to combine the above features with a 208-dimensional vector. The Extra Trees classifier used in this research is to classify PD datasets with a class-balanced algorithm distribution after SMOTE was used.

III. Results and Discussion

In this study, a new model for classifying the Parkinson's disease-dataset with class-imbalance data distribution based on voice signal is presented. The hold-out approach of 80% of the dataset used for training and 20% of it been used as test set, the 10-fold cross validation were used for the train and evaluation of the Extra Trees Classifier. The classification accuracy was used for determining the suggested approach as a performance measure. Table 1 shows the findings of the Extra Trees Classifier in PD Dataset classification and in the 10-fold cross-validation after SMOTE method were applied on the dataset. As seen in Table 1, Extra-tree classification findings of the PD classification with the ExtraTrees classifier were obtained in the PD dataset. First, the SMOTE was applied on the dataset. Secondly, ExtraTrees Classifier was used on the enhanced dataset; and finally, 10-fold cross-validation process was applied, which is repeated 10 times (the folds), with each of the 10 subsamples used exactly once as the validation data. Both methods, SMOTE and Extra Trees methods, are explained in details in the previous section. Using SMOTE and ExtraTrees methods exhibited encouraging findings in differentiating between healthy and PD patients based on their voice signals as presented in Table 1.

TABLE I
RESULT OF OUR PROPOSED MODEL COMPARED TO OTHER METHODS IN PARKINSON DETECTION.

Techniques	Accuracy
TQWT + MFCC + Concat [34]	86.90%
XGBoost + mRMR feature selection [35]	95.39%
Support Vector Machine + Wrappers Feature Subset Selection [36]	94.70%
mRMR + SVM-RBF [29]	86.00%
SMOTE and Random Forests [28]	94.89%
Our Proposed Model (SMOTE and ExtraTrees Classifier)	96.52%

Our proposed model using SMOTE and ExtraTrees Classifier has achieved the best accuracy (96.52%) with respect to the literature's classification accuracy,as shown in Table 1. Our proposed model has achieved higher classification accuracy than the used approaches in the related works. This work's obtained accuracy was 9.62% higher than the reported accuracy in [34] that used (TQWT + MFCC + Concat) approach. Also, it is 1.13% higher than the used approach in [35] ,which used XGBoost classifier with a minimum redundancy-maximum relevance (mRMR) feature selection. Moreover, our proposed approach has achieved accuracy 1.82% higher than the achieved accuracy in [36] that used Support Vector Machine classifier with a feature selection method called Wrappers Feature Subset Selection in [29], mRMR feature selection was used with SVM-RBF as a classifier. However, our proposed approach has achieved accuracy 10.86 % higher than their reported accuracy.

Although our proposed approach has achieved accuracy higher than the achieved accuracy in [28] where SMOTE technique with Random Forests Classifier were used. It is important to state that the obtained result showed imbalanced dataset and it proves that the balancing in the dataset is vital, especially in medical classification. Our proposed approach shows the capabilities of using machine learning in the medical field significantly.

TABLE II
RESULT OF OUR PROPOSED MODEL COMPARED TO OTHER METHODS IN PARKINSON DETECTION.

	Precision	Recall	F1-Score	Accuracy
Not-Parkinson	0.95	0.97	0.96	96.52%
Parkinson	0.96	0.94	0.95	

Table II shows the results obtained by our proposed model and measured by Precision, Recall, F1-Score and Support for the two classes in the dataset. This validate the accuracy which is mentioned in Table I and Table II.

IV. Conclusions

This paper describes the machine learning method's findings to automatically identify clinical data and aid in diagnosing patients with Parkinson's disease. Our proposed approach provides more consistent results among all the experiments that have been achieved in the related works. The impact of Optimal function selection, either automated by PCA or manually carried out, is clearly still being studied, and plays an essential role in improving the performance of machine learning. Moreover, a comparison of the different segments according to the way the vocal tract is articulated or restricted must be dealt with in future research on automatically detecting PD patients' voice. The precision of these approaches shows that these systems may be applied in the clinical field in addition to other clinical observations and tests to support diagnosis of PD, as a part of a multimodal system or additional to other tests and clinical observations.

Another future research involves the study of new multimodal systems, in order to support diagnosis, which combines speech with other inputs. In addition, one more move is to study this and other related structures in clinical environments to show their effectiveness thorough real clinical trials. Finally, the findings indicate that PD impacts the whole joint series and more specifically regulates phonetic units that involve a higher narrowing of the vocal system. For this function, it is possible to determine the existence of PD from speech by means of automated detectors by evaluating phonically controlled speech activities.

REFERENCES

[1] J. M. Bassetto, B. S. Zeigelboim, A. L. Jurkiewicz, and K. F. Klagenberg, "Neurotological findings in patients with Parkinson's disease," *Brazilian journal of otorhinolaryngology,* vol. 74, no. 3, pp. 350-355, 2008.

[2] B. Pillon, B. Dubois, G. Cusimano, A.-M. Bonnet, F. Lhermitte, and Y. Agid, "Does cognitive impairment in Parkinson's disease result from non-dopaminergic lesions?," *Journal of Neurology, Neurosurgery & Psychiatry,* vol. 52, no. 2, pp. 201-206, 1989.

[3] B. R. Bloem and M. Munneke, "Revolutionising management of chronic disease: the ParkinsonNet approach," *Bmj,* vol. 348, p. g1838, 2014.

[4] H.-L. Chen, G. Wang, C. Ma, Z.-N. Cai, W.-B. Liu, and S.-J. Wang, "An efficient hybrid kernel extreme learning machine approach for early diagnosis of Parkinson′ s disease," *Neurocomputing,* vol. 184, pp. 131-144, 2016.

[5] D. Calne, B. Snow, and C. Lee, "Criteria for diagnosing Parkinson's disease," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society,* vol. 32, no. S1, pp. S125-S127, 1992.

[6] H.-L. Chen *et al.*, "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach," *Expert systems with applications,* vol. 40, no. 1, pp. 263-271, 2013.

[7] C. Sommer and D. W. Gerlich, "Machine learning in cell biology–teaching computers to recognize phenotypes," *J Cell Sci,* vol. 126, no. 24, pp. 5529-5539, 2013.

[8] M. Nilashi, O. Bin Ibrahim, A. Mardani, A. Ahani, and A. Jusoh, "A soft computing approach for diabetes disease classification," *Health Informatics Journal,* vol. 24, no. 4, pp. 379-393, 2018.

[9] S. Ozer *et al.*, "Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI," *Medical physics,* vol. 37, no. 4, pp. 1873-1883, 2010.

[10] B. Fritzke, "Growing cell structures—a self-organizing network for unsupervised and supervised learning," *Neural networks,* vol. 7, no. 9, pp. 1441-1460, 1994.

[11] Ö. Eskidere, F. Ertaş, and C. Hanilçi, "A comparison of regression methods for remote tracking of Parkinson's disease progression," *Expert Systems with Applications,* vol. 39, no. 5, pp. 5523-5528, 2012.

[12] M. Hariharan, K. Polat, and R. Sindhu, "A new hybrid intelligent system for accurate detection of Parkinson's disease," *Computer methods and programs in biomedicine,* vol. 113, no. 3, pp. 904-913, 2014.

[13] T. Peterek, P. Dohnálek, P. Gajdoš, and M. Šmondrk, "Performance evaluation of Random Forest regression model in tracking Parkinson's disease progress," in *13th International Conference on Hybrid Intelligent Systems (HIS 2013)*, 2013: IEEE, pp. 83-87.

[14] D.-C. Li, C.-W. Liu, and S. C. Hu, "A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets," *Artificial Intelligence in Medicine,* vol. 52, no. 1, pp. 45-52, 2011.

[15] G. S. Babu and S. Suresh, "Parkinson's disease prediction using gene expression–A projection based learning meta-cognitive neural classifier approach," *Expert Systems with Applications,* vol. 40, no. 5, pp. 1519-1529, 2013.

[16] P.-F. Guo, P. Bhattacharya, and N. Kharma, "Advances in detecting Parkinson's disease," in *International Conference on Medical Biometrics*, 2010: Springer, pp. 306-314.

[17] F. Åström and R. Koker, "A parallel neural network approach to prediction of Parkinson's Disease," *Expert systems with applications,* vol. 38, no. 10, pp. 12470-12474, 2011.

[18] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, and M. Farahmand, "A hybrid intelligent system for the prediction of Parkinson's Disease progression using machine learning techniques," *Biocybernetics and Biomedical Engineering,* vol. 38, no. 1, pp. 1-15, 2018.

[19] R. Das, "A comparison of multiple classification methods for diagnosis of Parkinson disease," *Expert Systems with Applications,* vol. 37, no. 2, pp. 1568-1572, 2010.

[20] W. Froelich, K. Wrobel, and P. Porwik, "Diagnosis of Parkinson's disease using speech samples and threshold-based classification," *Journal of Medical Imaging and Health Informatics,* vol. 5, no. 6, pp. 1358-1363, 2015.

[21] D. Avci and A. Dogantekin, "An expert diagnosis system for parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine," *Parkinson's disease,* vol. 2016, 2016.

[22] A. Ozcift, "SVM feature selection based rotation forest ensemble classifiers to improve computer-aided diagnosis of Parkinson disease," *Journal of medical systems,* vol. 36, no. 4, pp. 2141-2147, 2012.

[23] W.-L. Zuo, Z.-Y. Wang, T. Liu, and H.-L. Chen, "Effective detection of Parkinson's disease using an adaptive fuzzy k-nearest neighbor approach," *Biomedical Signal Processing and Control,* vol. 8, no. 4, pp. 364-373, 2013.

[24] J. A. Logemann, H. B. Fisher, B. Boshes, and E. R. Blonsky, "Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients," *Journal of Speech and hearing Disorders,* vol. 43, no. 1, pp. 47-57, 1978.

[25] L. Parisi, N. RaviChandran, and M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Systems with Applications,* vol. 110, pp. 182-190, 2018.

[26] M. S. Islam, I. Parvez, H. Deng, and P. Goswami, "Performance comparison of heterogeneous classifiers for detection of Parkinson's disease using voice disorder (dysphonia)," in *2014 International Conference on Informatics, Electronics & Vision (ICIEV)*, 2014: IEEE, pp. 1-7.

[27] V. Despotovic, T. Skovranek, and C. Schommer, "Speech Based Estimation of Parkinson's Disease Using Gaussian Processes and Automatic Relevance Determination," *Neurocomputing,* 2020.

[28] K. Polat, "A Hybrid Approach to Parkinson Disease Classification Using Speech Signal: The Combination of SMOTE and Random Forests," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, 2019: IEEE, pp. 1-3.

[29] C. O. Sakar *et al.*, "A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform," *Applied Soft Computing,* vol. 74, pp. 255-263, 2019.

[30] [Online]. Available: https://archive.ics.uci.edu/ml/datasets.html, (last accessed: April, 2020). .

[31] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research,* vol. 16, pp. 321-357, 2002.

[32] Y. Xie, Y. Liu, and Q. Fu, "Imbalanced data sets classification based on SVM for sand-dust storm warning," *Discrete Dynamics in Nature and Society,* vol. 2015, 2015.

[33] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Machine learning,* vol. 63, no. 1, pp. 3-42, 2006.

[34] H. Gunduz, "Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets," *IEEE Access,* vol. 7, pp. 115540-115551, 2019.

[35] I. Nissar, D. Rizvi, S. Masood, and A. Mir, "Voice-Based Detection of Parkinson's Disease through Ensemble Machine Learning Approach: A Performance Study," *EAI Endorsed Transactions on Pervasive Health and Technology,* vol. 5, no. 19, 2019.

[36] G. Solana-Lavalle, J.-C. Galán-Hernández, and R. Rosas-Romero, "Automatic Parkinson disease detection at early stages as a pre-diagnosis tool by using classifiers and a small set of vocal features," *Biocybernetics and Biomedical Engineering,* vol. 40, no. 1, pp. 505-516, 2020.