Mathematical Biosciences and Engineering

MBE, 20(5): 8085–8102.
DOI: 10.3934/mbe.2023351
Received: 11 January 2023
Revised: 08 February 2023
Accepted: 13 February 2023
Published: 24 February 2023

http://www.aimspress.com/journal/MBE

Research article

Intelligent personalized diagnosis modeling in advanced medical system for Parkinson's disease using voice signals

Pengcheng Wen[1], Yuhan Zhang[2,*] and Guihua Wen[3]

[1] College of Intelligent Systems Science and Engineering, Hubei University for Nationalities, Enshi 445000, China
[2] Southern Medical University, Affiliated Dongguan Songshan Lake Central Hospital, Dongguan 523000, China
[3] School of Computer Science & Engineering, South China University of Technology, Guangzhou 510000, China

* **Correspondence:** Email: zyhan2002@126.com.

Abstract: Currently, machine learning methods have been utilized to realize the early detection of Parkinson's disease (PD) by using voice signals. Because the vocal system of each person is unique, and the same person's pronunciation can be different at different times, the training samples used in machine learning become very different from the speech signal of the patient to be diagnosed, frequently resulting in poor diagnostic performance. On this account, this paper presents a new intelligent personalized diagnosis method (PDM) for Parkinson's disease. The method was designed to begin with constructing new training data by assigning the best classifier to each training sample composed of features from the speech signals of patients. Subsequently, a meta-classifier was trained on the new training data. Finally, for the signal of each test patient, the method used the meta-classifier to select the most appropriate classifier, followed by adopting the selected classifier to classify the signal so that the more accurate diagnosis result of the test patient can be obtained. The novelty of the proposed method is that the proposed method uses different classifiers to perform the diagnosis of PD for diversified patients, whereas the current method uses the same classifier to diagnose all patients to be tested. Results of a large number of experiments show that PDM not only improves the performance but also exceeds the existing methods in speed.

Keywords: Parkinson's disease; personalized diagnosis method (PDM); machine learning; classification; voice signals

1. Introduction

Parkinson's disease (PD) affects a large number of people's lives for a long time by social isolation and financial burdens, for it is not easily cured [1]. Although drugs can significantly alleviate symptoms of the disease at early stages, it is very difficult to evaluate the severity of PD accurately. Moreover, it usually takes several days to diagnose the patient through the complete evaluation of symptoms, leading to the decrease of diagnosis accuracy. In addition, it is inconvenient to monitor patients with PD in the hospital, consuming more medical resources and increasing medical costs. With such a background, it is necessary to develop an intelligent device to facilitate the diagnosis and prediction of PD anytime and anywhere [2]. Once a person has the risk for PD, he then goes to the hospital for the diagnosis and treatment. Lots of symptoms of PD can be observed at early stages through physical signals [3,4], including voice [5,6], gait [7,8], handwriting [9], electroencephalography (EEG) [10], magnetic resonance imaging (MRI) [11] and facial expressions [12], apart from the multi-models in consideration [13]. These signals can be captured by lots of sensors and analyzed by machine learning methods automatically [14,15]. Among them, the voice signals are very significant for the early diagnosis [16], for the vocal impairment of a patient is easily identified at the beginning and generally continues many years before the diagnosis of PD is determined eventually. Patients often show a lot of vocal symptoms such as disfluencies and monotonous voices, as well as inaccurate consonant articulation [17]. In addition, the measurement method of PD is noninvasive, simple and easy to use [18]. As a result, the voice signals have been widely applied to recognize and track symptoms of PD through speech features and machine learning methods [19,20]. These features are extracted by feature engineering to improve the diagnosis accuracy [21–23] and afterward refined by using feature selection methods like filter and wrapper [24], popular relief feature selection method and particle swarm optimization algorithm. Based on the determined features, many machine learning methods are applicable [25], including neural networks (NN), softmax, naive Bayes classifier (NBC), support vector machine (SVM), decision trees (DT), linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA). It has been demonstrated that PD is able to be well predicted by the ensemble methods [26] like random forest (RF), AdaBoost and bagging methods [27]. Furthermore, these machine learning methods can be integrated [28], while feature selection and classification mechanisms can also be combined [29–33].

Although deep neural networks have been frequently utilized for the diagnosis of PD with better results [34,35], they work on the condition that the captured voice signals have a larger size [36]. For example, convolutional neural networks can learn features from the raw data, doing well in such situations. However, when there are not enough training data available, they become less effective, which often occurs in medical fields [37]. This case very much applies to the diagnosis of PD, in which the labeled samples are insufficient. This is due to the fact that specialized knowledge is required to label samples, costing a lot of labor and time. Another serious problem is the disaster of dimension, as the voice signals are high-dimensional [38]. Theoretically, as the number of features increases, the number of speech signals required to produce a reliable statistical conclusion increases exponentially. However, it is hard to obtain a large number of speech signals in practice. Considering the above situation, it

is a great challenge to create nice predictive models using machine learning methods when there is little training data but with many features. In biomedical settings, this is a typical scenario [39], in which both feature selection [40] and the transfer learning [41,42] have been tried to solve the problem, yet the effects are limited. Specifically, the deep learning methods usually cannot obtain the ideal effect on the limited training data. In such a circumstance, extracting features through feature engineering and designing a better machine learning diagnosis method for PD is still the best choice. The existent machine learning methods for PD are generally trained to obtain the model. Subsequently, the same model is applied to classify test samples, easily leading to poor performance in diagnosis. This is inconsistent with the fact that humans dynamically change their methods based on the current test samples [43]. This is because the vocal system of each person is different at different times, and the training speech signals used in machine learning may be very different from the patient's speech in the clinical diagnosis [44], resulting in the inaccurate diagnosis. Among the other reasons are the subject gender, the voice environment and patient's condition. Furthermore, the captured training data may be different from those in the clinical diagnosis. Thus, the diagnosis abilities of machine learning methods are different even when applied to the same test sample. It is however a fact that clinical diagnosis and machine-based diagnosis are complementary to each other. It has been discovered that a classifier works well for some test samples, but it erroneously classifies the others. Particularly, when a pair of classifiers are utilized to categorize test speech signals, their classification results may become completely opposite. In this situation, it is reasonable to dynamically select the suitable classifier for the given test patient with PD so as to perform the personalized diagnosis. In this paper, a new personalized diagnosis method (PDM) for Parkinson's disease is proposed. For the voice signals of a given patient, PDM first uses a Bayesian learning method to analyze the situation of the patient from the global perspective and then selects the most appropriate diagnosis method to diagnose the patient with PD. The proposed method is different from the currently used methods for PD in that it directly endows each training signal with the classifier name as labels to train a meta-classifier, which in turn is used to select a suitable diagnosis method for testing patients.

2. Methods

PDM was trained using lots of voice signals of the previously diagnosed patients and healthy people, and it was then applied to predict the test patients by speech signals.

2.1. Framework of Parkinson's disease diagnosis

PD easily leads to muscle deformation and failure to create distinguished pronunciation, leading to the variation of speech signals. Based on this point, the features of a patient's natural speech signals can be utilized to monitor the developmental progress of PD. The general framework is shown in Figure 1, involving the data acquisition for the voice signals, preprocessing method, feature extraction and the classifier for the classification of PD. When the voice signals of the patient are captured, they are used to extract features, and then the disease is predicted using the machine learning technique. This paper emphasizes designing the machine learning method suitable for predicting PD disease, based on the large number of the voice signal samples.

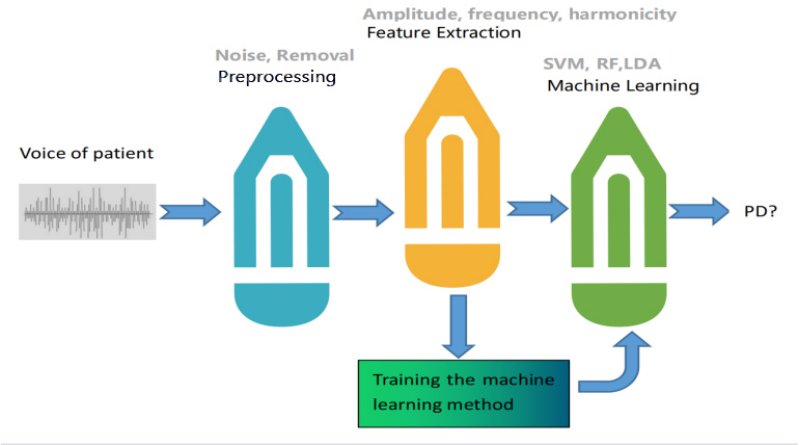

Figure 1. Framework of Parkinson's disease diagnosis.

2.1.1. Voice signal preprocessing

Voice tasks refer to the method of collecting patient voice signals. The voice signals can be obtained through the dialogue between the doctor and the patient as well as the given text that a patient reads, by virtue of various devices like smart phones [45]. Subsequently, these signals are processed, including removing the noise outside the effective frequency range for PD. As the maximum amplitude is not useful for PD and may vary from person to person, the voice signal is normalized. Furthermore, more complex preprocessing methods can be used to separate voiced signals and unvoiced signals as well as to remove outliers, in order to improve the quality of voice signals. In addition, the machine learning method needs a large number of such processed voice signals which are captured from the clinical data, and each signal is annotated with a tag by doctors. The tag is one when the speech signal is from patients diagnosed with PD. The tag is zero for the voice signal of a healthy person. Due to the large amount of training data being hard to prepare manually, any machine learning method is prone to over-fitting. In such a case, the data augmentation method can be used to increase the number of voice signals, which has been validated in the previous work.

2.1.2. Feature extraction

Features of each voice signal are extracted as the input of the machine learning method to perform the diagnosis of PD. These features can be handcrafted or learned from a large number of voice signals through deep learning methods. In the previous studies, most of the features are handcrafted, including fundamental frequency, jitter, shimmer, Mel-Frequency Cepstral Coefficients, etc. These features are then applied to identify the cases of PD when they are combined with an advanced classifier. However, these features heavily rely on the human experience and professional knowledge. Due to the uncertainty of the human experience and the limitation of professional knowledge, the designed features are often incomplete and even contradictory with each other. The ideal features should be relevant to the classification problem, while the relationship between these features should be irrelevant and orthogonal in the feature space. The features that can be combined from other features are redundant, such that they should be removed. Simultaneously, these features should be consistent

to ensure that the noisy features have been removed. In such circumstances, efficient feature selection methods should be applied to select the suitable ones.

Apart from the above, since deep neural networks are most beneficial to deal with images, the voice signal of patients can be transformed to image data using short-time Fourier transform [46], in which both time and frequency information of signals are preserved. Subsequently, deep neural networks can be applied to extract features from the image of speech. However, these methods require a larger number of training data. When few training data are available, their extracted features easily make the model over-fit the training data, leading to bad generalization ability. It is much expected that the model can automatically learn features well even if there is only a small number of training samples. As a result, transfer learning can be used to solve this issue by obtaining features from other domains [47].

2.1.3. Machine learning method for classification

After features of the speech signals were extracted, a couple of machine learning approaches were selected and then trained to identify PD. Generally, the best machine learning method should be selected, which is most useful to classify an individual person into healthy or a patient with PD. From previous research, there are various machine learning methods, such as NN, RF and SVM. Although the performance of the machine learning method might vary with the selected data sets and extracted features, RF and SVM were proven to provide the optimal results [48], while NBC performs well with reduced features [49]. However, when classes cannot be separated linearly in the feature space, a non-linear mapping function should be applied to transform features to the higher dimensional space, e.g., by the kernel techniques.

2.2. PDM

In practice, various doctors diagnose the same person in different ways, leading to different results, i.e., partly correct and partly wrong. In such case, the best doctor should be selected to diagnose the person. However, the existent methods for PD are designed such that one machine learning method was trained and then used to classify all persons, which easily leads to some incorrect classification. To address the dilemma, PDM was proposed for Parkinson's disease, as illustrated in Figure 2. This method was different from the ensemble technique for PD. In principle, the latter averaged all ensemble classifiers so that it was more stable to identify patients with PD. However, the ensemble technique reduced the ability of the best classifier in the ensemble. Theoretically, it is inferior to the best classifier. Additionally, it is unlike the model selection technique. The latter seeks the best classifier across all test samples as opposed to only one at a time. Because each doctor has unique strengths, it is believed that each patient has a particular ideal doctor to identify PD, where the progress of PD may be different from person to person. Thus, PDM determines the most appropriate classifier for each test patient, instead of using the same classifier to classify all test patients with PD. In our method, a new training set, $D_c = \{(x_i, c_i)\}$, should be created, where x_i is the training sample composed of features from the patient speech signal, and c_I is the tag as the best classifier to classify x_i. Subsequently, a meta-classifier is trained on this new training set and then used to choose the best classifier for each test patient.

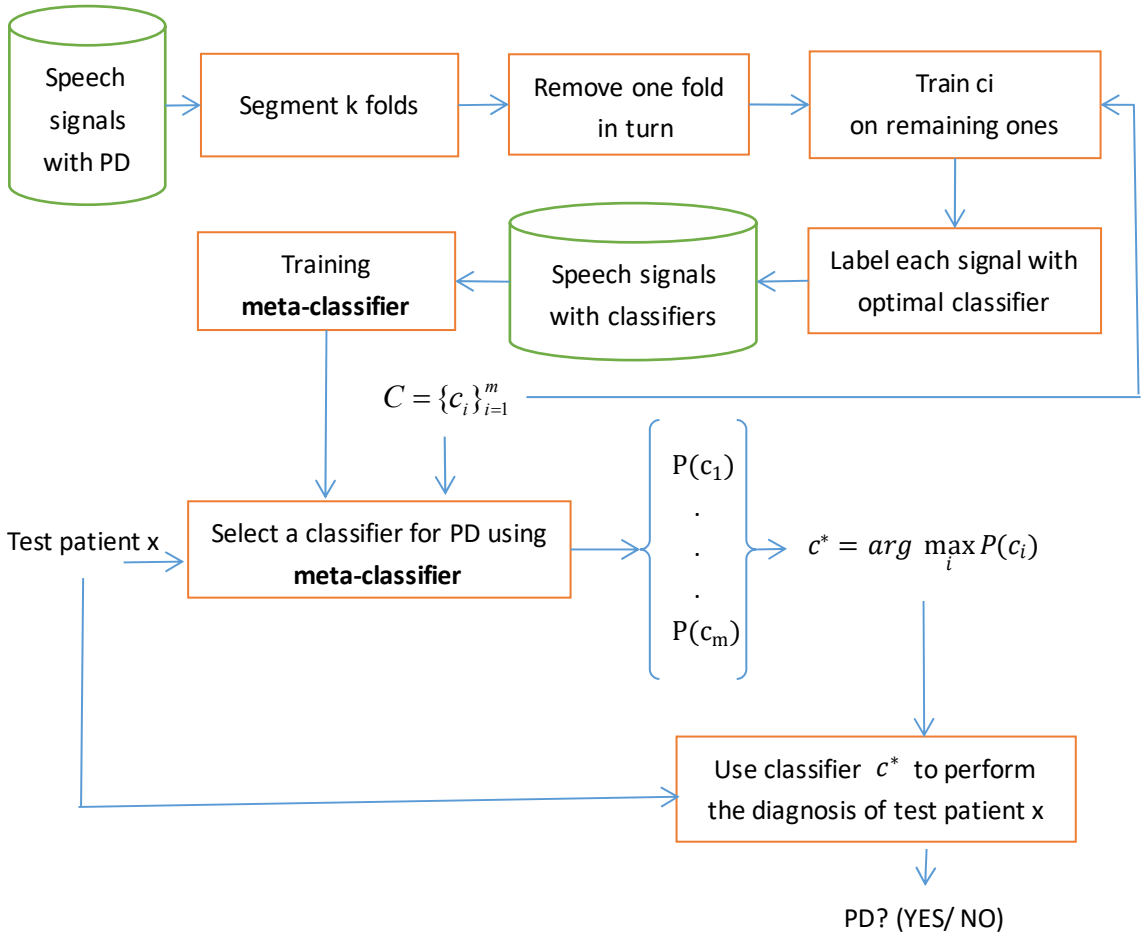

Figure 2. Flowchart of PDM for Parkinson's disease on the voice signal.

2.2.1. Create classifiers

PDM used lots of candidate classifiers to simulate the different doctors. It first selected the optimal classifier from the candidate ones through patients' speech signals and then used the system to perform the diagnosis of PD. Thus, all candidate classifiers should be powerful and complementary, so as to form a diversity to correctly classify all kinds of testing patients. To date, there have been many methods available to generate candidate classifiers. Therefore, we can directly adopt the method of generating base classifiers in ensemble learning. For example, for the same machine learning method, classifiers with different performances can be created by selecting different training subsets. Particle swarm optimization method can be used to select an optimized training subset for each candidate classifier that is optimal for identifying PD. The advantages are that a large number of candidate classifiers can be generated easily. However, these created classifiers may become easily similar, resulting in poor diversity. The other method uses heterogeneous machine learning methods to create candidate classifiers, with its superiority in maintaining the diversity easily. Our method used this method to create the candidate classifiers to perform the diagnosis of PD.

2.2.2. Labeling each voice sample with the classifier name

Let $X = \{x_i | x_i \in R^n)\}$ be the training data composed of n-dimensional feature vectors from speech signals of both patients with PD and healthy persons, $Y = \{y_i | y_i \in \{0,1\}\}$ be the matching label set and $C = \{c_i\}$ be composed of classifiers. The classifier c_i can be used to categorize each voice signal feature vector x_i and then measure the probability that x_i can be correctly classified by c_i. We use the k-fold cross validation method to determine the training subset from the whole data. That is, feature vectors in (k−1) folds are applied to form the training subset, on which each classifier is trained to classify all data. If the feature vector is classified correctly, this classifier is suitable to perform the diagnosis of PD, so that this feature vector is labeled by the name of the classifier. According to the k-fold cross-validation method, the whole data are decomposed as follows:

$$X = \cup_{j=1}^{k} X_j , X_i \cap X_j = \phi, |X_i| = |X_j| \tag{1}$$

$$Y = \cup_{j=1}^{k} Y_j , |Y_j| = |X_j| \tag{2}$$

Suppose that the classifier c_i is trained on the subset X_i, and its decision function is defined by

$$f_{c_i, X_i \subset X}: X_i \rightarrow Y_i \tag{3}$$

The prior probability that the classifier c_i can correctly classify any patient in the training data can be computed as follows, where $1|f_{c,X-Xi}(x_m) == y_m$ indicates that if x_m is correctly classified, the value is 1, and else it is zero.

$$P(c_i) = \frac{1}{k} \sum_{j=1}^{k} \sum_{m=1}^{|X|} 1 \mid f_{c,X-X_j}(x_m) == y_m \tag{4}$$

Similarly, the prior probability that the classifier c_i can correctly perform the diagnosis of PD for x_m is computed by

$$P(x_m \mid c_i) = \frac{1}{k} \sum_{j=1}^{k} 1 \mid f_{c_i,X-X_j}(x_m) == y_m \tag{5}$$

However, our goal is to choose the optimal classifier for the test patient. That is, $P(c_i|x_m)$ is required, which is the probability that each classifier should be chosen for the test patient. Subsequently, the classifier with the biggest probability is selected to classify each test patient. $P(c_i|x_m)$ can be computed according to the Bayesian theorem:

$$P(c_i|x_m) = \frac{P(x_m|c_i)P(c_i)}{P(x_m)} \tag{6}$$

As the test patient x_m is the same for all classifiers, $P(x_m)$ can be omitted. This is consistent with the assumption that NBC makes. Subsequently, the classifier name is tagged to each training speech signal so as to create a new training data set D, where the probability of the selected classifier is bigger than the given threshold θ.

$$D_i = \{(x_m, c_i) \mid x_m \in X, c_i \in C, P(c_i \mid x_m) > \theta\} \tag{7}$$

$$S = \bigcup_{i=1}^{|C|} D_i \tag{8}$$

$$D = \{(x, c) \mid x \in S, c = \arg \max_j P(c_j \mid x)\tag{9}$$

2.2.3. Selecting the classifier for the test sample

Once D is constructed, with samples labeled with the classifier names, a meta-classifier, ϕ, is trained on D so as to obtain a new decision function:

$$h_{\varphi,D}: X \to 2^C \tag{10}$$

$$c = \arg \max_i P(c_i \in h_{\varphi,D}(x)) \tag{11}$$

Subsequently, the classifier c selected by the meta-classifier can be applied to perform the diagnosis for the test patient x.

3. Experiments and results

Experiments on two benchmark data sets were conducted to validate the proposed system by using the most used classifiers. Despite the fact that the validation can use a variety of evaluation indicators, the accuracy was often taken as the indicator, the indicator in our experiments. Additionally, the confusion matrix was used.

3.1. Data sets

Parkinson's Disease Classification Data Set (D1) [50]. The data were composed of two classes, 754 features and 756 samples, among which 192 were from patients with PD. The sets' features included features of time frequency, vocal fold features, features of tunable Q-factor wavelet transform (TQWT), Mel-Frequency Cepstral Coefficients, features by wavelet transform and so on.

Oxford Parkinson's Disease Detection Data set (D2) [51]. The data include the unified Parkinson's disease rating scale (UPDRS) for evaluating the symptoms severity of the patients with PD. In order to perform the classification, we constructed new data with two classes from this data. One class was composed of patients with the symptom severity less than 0.5, having 3154 samples, and the remaining 2721 samples constituted the other class. As a result, the newly constructed data had two classes and 5875 samples with 18 features.

3.2. Comparison of different classifiers for PD

In previous work, lots of classifiers have been applied to recognize patients with PD. In this section, some experiments were conducted to validate these classifiers for the recognition of PD and then choose candidate classifiers from them for PDM, in which all classifiers use their default

parameters. As shown in Tables 1 and 2, RF obtained the best performance on two data sets and also had the smaller standard deviation, indicating that it is very stable. This is consistent with the previous research. Furthermore, RF significantly outperforms DT, because it is composed of lots of DT. By comparison, SVM and NN worked better on different data, respectively. Simultaneously, SVM generally worked better on the limited data, so it was taken as the meta-classifier in PDM that arranged the classifier for each patient to be tested.

Table 1. Experimental results of different classifiers on D1.

Method	Accuracy
Softmax	0.8651 ± 0.0112
LDA	0.6574 ± 0.0365
QDA	0.4500 ± 0.2486
NBC	0.7460 ± 0.0034
DT	0.8174 ± 0.0307
RF	0.8809 ± 0.0172
AdaBoost	0.7923 ± 0.0388
SVM	0.7434 ± 0.0020
NN	0.8598 ± 0.0340

Table 2. Experimental results of different classifiers on D2.

Methods	Accuracy
Softmax	0.6306 ± 0.0074
LDA	0.6363 ± 0.0049
QDA	0.6106 ± 0.0141
NBC	0.6165 ± 0.0212
DT	0.9154 ± 0.0062
RF	0.9260 ± 0.0079
AdaBoost	0.6381 ± 0.0107
SVM	0.8945 ± 0.0021
NN	0.6929 ± 0.0098

3.3. Comparison of proposed method with different candidate classifiers for the performance

As demonstrated above, RF worked best among all classifiers. DT is also nice but simpler. DT has been widely used for the classification, as it can be constructed easily with higher efficiency. It is readable and descriptive, so it is very useful for the diagnosis of PD. In practice, DT also worked fast, as the maximum calculation times of each prediction did not exceed the depth of the decision tree. RF was a method that used the concept of ensemble learning to integrate various decision trees, handling a large number of input features without variable deletion and working effectively on the larger training data. Also, it could estimate the importance of variables for the classification. Furthermore, the

accuracy of RT kept constant when lots of data were missing. Accordingly, both DT and RF were selected as candidate classifiers for PDM. They were compared via experiments so as to validate the proposed PDM. As shown in Table 3, PDM outperforms both RF and DT in accuracy on any data measurement. Although RF was composed of lots of DT, DT was a good complement to RF, classifying the same patient by virtue of different abilities.

Table 3. Experimental results of PDM on Parkinson's disease data sets.

Database	Method	Accuracy
D1	RF	0.8678 ± 0.0283
	DT	0.7910 ± 0.0353
	PDM	0.8770 ± 0.0334
D2	RF	0.9232 ± 0.0111
	DT	0.9244 ± 0.0100
	PDM	0.9275 ± 0.0109

However, due to the imbalance of data, for the performance analysis of a classifier, accuracy is a poor performance indicator. A classifier that has been trained with the unbalanced data will typically predict a class that belongs to the majority and disregard the minority. In such a case, the confusion matrix can be used to make the comparison in a more fair way. In the field of machine learning, the confusion matrix is a specific matrix used to present the visualization effect of classification performance, whose columns represent the predicted values and rows represent the actual classes. By using the confusion matrix, we can easily indicate whether different classes are confused. The confusion matrix, which is a two-row and two-column table in the prediction analysis, is made up of false positives, false negatives, true positives, and true negatives. The matrix allowed us to do more analysis such as the precision and recall, not just limited to the accuracy. For example, if there were 564 healthy samples but only 192 samples with Parkinson's disease in a data set, some classifiers may be more likely to predict all samples as healthy ones. As illustrated in Tables 4 to 9, PDM not only outperformed RF and DT by accuracy on the two data but also obtained better results from the viewpoint of the confusion matrix. In a word, PDM classifies the patients with PD more accurately.

Table 4. Confusion matrix of DT on data D1.

predicted→	Parkinson	~Parkinson
Parkinson	108	84
~Parkinson	74	490

Table 5. Confusion matrix of RF on data D1.

predicted→	Parkinson	∼Parkinson
Parkinson	107	85
∼Parkinson	15	549

Table 6. Confusion matrix of PDM on data D1.

predicted→	Parkinson	∼Parkinson
Parkinson	116	76
∼Parkinson	17	547

Table 7. Confusion matrix of DT on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2914	240
∼Parkinson	211	2510

Table 8. Confusion matrix of RF on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2927	227
∼Parkinson	217	2504

Table 9. Confusion matrix of PDM on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2939	215
∼Parkinson	211	2510

3.4. Comparison of proposed method with different candidate classifiers for the applications

Lots of systems have been applied to identify PD at the early stage, reducing the uncomfortable clinical check and the strain of doctors. The typical system is a smartphone app, which uses multi-modal fusion to supply useful diagnosis recommendations for PD patients. The system captures speech signals from sensors installed on the smartphone [52]. In such applications, the speed and simplicity of machine learning methods for PD become more important. Among all compared classifiers, NBC and DT have been validated in recognizing PD in speed. They can be easily installed in the smartphone. In this section, they are selected as candidate classifiers for experiments. NBC can achieve good results even in many complex situations, despite its simple ideas and simplistic assumptions. NN have been applied to recognize PD [53], generally referring to the simple feed forward neural network. It has strong nonlinear mapping ability. Thus, NBC, DT, and NN were selected as candidate classifiers for PDM and then compared through experiments. Table 10 shows that PDM outperformed all compared methods on any data in terms of the accuracy and the standard deviation. It is also illustrated that DT is really complementary to NBC and NN, for they had different abilities to classify voice samples of PD. Simultaneously, PDM was more stable, as it obtained a lesser standard deviation.

Table 10. Experimental results of PDM on Parkinson's disease.

Database	Methods	Accuracy
D1	NBC	0.7447 ± 0.0072
	DT	0.8121 ± 0.0286
	PDM	0.8703 ± 0.0156
D2	NN	0.6917 ± 0.0069
	DT	0.9220 ± 0.0070
	PDM	0.9220 ± 0.0059

In addition, these methods were compared by means of the confusion matrix. Table 11 to Table 13 reveal that NBC worked much worse on data D1. Although its overall accuracy was 0.7447 with the standard deviation 0.0072, it classified almost all samples into the healthy persons, including patients with Parkinson's disease. That is, it failed to recognize PD correctly. By comparison, PDM worked best on the whole, surpassing both NBC and DT on data D1. However, it was still influenced by NBC, leading to some classification errors. This illustrates that PDM heavily depended on its candidate classifiers, which should be selected carefully. Simultaneously, Tables 14 to 16 suggest that NN did not work well on data D2. Although PDM and DT obtained the same accuracy, PDM obtained the better results from the viewpoint of the confusion matrix and the smaller standard deviation. It is more stable than DT on D2.

Table 11. Confusion matrix of NBC on D1.

predicted→	Parkinson	∼Parkinson
Parkinson	1	191
∼Parkinson	2	562

Table 12. Confusion matrix of DT on D1.

predicted→	Parkinson	∼Parkinson
Parkinson	126	66
∼Parkinson	76	488

Table 13. Confusion matrix of PDM on D1.

predicted→	Parkinson	∼Parkinson
Parkinson	103	89
∼Parkinson	9	555

Table 14. confusion matrix of NN on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2492	662
∼Parkinson	1149	1572

Table 15. Confusion matrix of DT on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2912	242
∼Parkinson	216	2505

Table 16. Confusion matrix of PDM on data D2.

predicted→	Parkinson	∼Parkinson
Parkinson	2923	231
∼Parkinson	227	2494

4. Conclusions

Parkinson's disease should be diagnosed and treated as early as possible in order to delay the process of neurodegeneration. This paper presents a new personalized diagnosis method for Parkinson's disease. The novelty of our method is that it can use different classifiers to perform the diagnosis of PD from patient to patient, whereas the current method always uses the same classifier to perform the diagnosis for all test patients. Particularly, the method to create a meta-classifier is also proposed, by which the appropriate classifier can be automatically selected for each test patient. The proposed method reveals its advantage in that it can select simple diagnostic methods as candidate ones, superior to some complicated methods such as ensemble methods in both performance and efficiency. This method is more suitable for running in portable devices such as smart phones, so that each person can diagnose himself any time and then treat himself as early as possible. In our method for PD, the candidate classifiers are very crucial, as they should be not only simple but also complementary. Simultaneously, this paper only considers the diagnosis method, without considering the feature extraction method. As a matter of fact, the feature extraction is highly related to the diagnosis method and hence should be considered at the same time. All these issues are to be further explored in the future.

Acknowledgments

The study was supported by Dongguan Key Projects of Social Science and Technology Development Plan Project (Grant No. 202050715024222) and the China National Science Foundation (Grant No. 62176095).

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. S. Singh, W. Xu, Robust detection of Parkinson's disease using harvested smartphone voice data: a telemedicine approach, *Telemed. e-Health*, **26** (2020), 327–334. https://doi.org/10.1089/tmj.2018.0271
2. A. Tsanas, M. Little, P. McSharry, L. Ramig, Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests, *Nat. Preced.*, **2009** (2009). https://doi.org/10.1038/npre.2009.3920.1

3. T. L. Yang, C. H. Lin, W. L. Chen, H. Y. Lin, C. S. Su, C. K. Liang, Hash transformation and machine learning-based decision-making classifier improved the accuracy rate of automated Parkinson's disease screening, *IEEE Trans. Neural Syst. Rehabil. Eng.*, **28** (2020), 72–82. https://doi.org/10.1109/TNSRE.2019.2950143
4. O. Y. Chen, F. Lipsmeier, H. Phan, J. Prince, K. I. Taylor, C. Gossens, et al., Building a machine-learning framework to remotely assess Parkinson's disease using smartphones, *IEEE Trans. Biomed. Eng.*, **67** (2020), 3491–3500. https://doi.org/10.1109/TBME.2020.2988942
5. T. Tuncer, S. Dogan, U. R. Acharya, Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels, *Biocybern. Biomed. Eng.*, **40** (2020), 211–220. https://doi.org/10.1016/j.bbe.2019.05.006
6. S. A. Mostafa, A. Mustapha, M. A. Mohammed, R. I. Hamed, N. Arunkumar, S. H. Khaleefah, et al., Examining multiple feature evaluation and classification methods for improving the diagnosis of Parkinson's disease, *Cognit. Syst. Res.*, **54** (2019), 90–99. https://doi.org/10.1016/j.cogsys.2018.12.004
7. I. EI. Maachi, G. Bilodeau, W. Bouachir, Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait, *Expert Syst. Appl.*, **143** (2020), 113075. https://doi.org/10.1016/j.eswa.2019.113075
8. S. Aich, P. M. Pradhan, S. Chakraborty, H. Kim, M. Joo, J. Park, et al., Design of a machine learning-assisted wearable accelerometer-based automated system for studying the effect of dopaminergic medicine on gait characteristics of Parkinson's patients, *J. Healthcare Eng.*, **2020** (2020), 1823268. https://doi.org/10.1155/2020/1823268
9. S. Rosenblum, S. Meyer, A. Richardson, S. Hassin-Baer, Patients' self-report and handwriting performance features as indicators for suspected mild cognitive impairment in Parkinson's disease, *Sensors*, **22** (2022). https://doi.org/10.3390/s22020569
10. Q. T. Ly, A. M. Ardi Handojoseno, M. Gilat, R. Chai, K. Martens, M. Georgiades, et al., Detection of turning freeze in Parkinson's disease based on S-transform decomposition of EEG signals, in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (2017), 3044–3047. https://doi.org/10.1109/EMBC.2017.8037499
11. Z. Y. Shu, S. J. Cui, X. Wu, P. Huang, P. P. Pang, Y. Xu, et al., Predicting the progression of Parkinson's disease using conventional MRI and machine learning: An application of radiomic biomarkers in whole-brain white matter, *Magn. Reson. Med.*, **85** (2021), 1611–1624. https://doi.org/10.1002/mrm.28522
12. L. Yang, X. Chen, J. Zhang, Q. Guo, J. Zhang, X. Zou, et al., Changes in facial expressions in patients with Parkinson's disease during the phonation test and their correlation with disease severity, *Comput. Speech Lang.*, **72** (2022). https://doi.org/10.1016/j.csl.2021.101286
13. J. Archila, A. Manzanera, F. Martinez, A multimodal Parkinson quantification by fusing eye and gait motion patterns, using covariance descriptors, from non-invasive computer vision, *Comput. Methods Programs Biomed.*, **215** (2022). https://doi.org/10.1016/j.cmpb.2021.106607
14. L. Gutierrez-Loaiza, W. Alfonso-Morales, Morpho-logical neural networks for Parkinson detection through speech signals, in *IEEE Colombian Conference on Applications of Computational Intelligence*, (2020). https://doi.org/10.1109/ColCACI50549.2020.9247918
15. M. G. Krokidis, G. N. Dimitrakopoulos, A. G. Vrahatis, C. Tzouvelekis, D. Drakoulis, T. P. Exarchos, et al., A sensor-based perspective in early-stage Parkinson' disease: current state and the need for machine learning processes, *Sensors*, **22** (2022). https://doi.org/10.3390/s22020409

16. A. S. Gullapalli, V. K. Mittal, Early detection of Parkinson's disease through speech features and machine learning: a review, in *ICT with Intelligent Applications*, Springer nature, (2022), 203–212. https://doi.org/10.1007/978-981-16-4177-0_22
17. R. Viswanathan, P. Khojasteh, B. Aliahmad, S. P. Arjunan, P. Kempster, K. Wong, et al., Efficiency of voice features based on consonant for detection of Parkinson's disease, in *2018 IEEE Life Sciences Conference*, 2018. https://doi.org/49-52.10.1109/LSC.2018.8572266
18. T. Khan, L. E. Lundgren, D. G. Anderson, I. Nowak, M. Dougherty, A. Verikas, et al., Assessing Parkinson's disease severity using speech analysis in non-native speakers, *Comput. Speech Lang.*, **61** (2020). https://doi.org/10.1016/j.csl.2019.101047
19. D. Gupta, A. Julka, S. Jain, T. Aggarwal, A. Khanna, N. Arunkumar, et al., Optimized cuttlefish algorithm for diagnosis of Parkinson's disease, *Cognit. Syst. Res.*, **52** (2018). https://doi.org/10.1016/j.cogsys.2018.06.006
20. M. Pramanik, R. Pradhan, P. Nandy, Biomarkers for detection of Parkinson's disease using machine learning-A short review, in *Soft Computing Techniques and Applications*, Springer nature, (2020), 461–475. https://doi.org/10.1007/978-981-15-7394-1_43
21. A. UI Haq, J. Li, M. H. Memon, J. Khan, A. Malik, A. Ali, et al., Feature selection based on L1-norm support vector machine and effective recognition system for Parkinson's disease using voice recordings, *IEEE Access*, **2019** (2019), 37718–37734. https://doi.org/10.1109/ACCESS.2019.2906350
22. S. Arora, L. Baghai-Ravary, A. Tsanas, Developing a large scale population screening tool for the assessment of Parkinson' disease using telephone-quality voice, *J. Acoust. Soc. Am.*, **145** (2019), 2871–2884. https://doi.org/10.1121/1.5100272
23. M. Nilashi, O. Ibrahim, S. Samad, H. Ahmadi, L. Shahmoradi, E. Akbari, An analytical method for measuring the Parkinson's disease progression: a case on a Parkinson's telemonitoring dataset, *Measurement*, **136** (2019), 545–557. https://doi.org/10.1016/j.measurement.2019.01.014
24. A. B. Soliman, M. Fares, M. M. Elhefnawi, M. Al-Hefnawy, Features selection for building an early diagnosis machine learning model for Parkinson's disease, in *2016 Third International Conference on Artificial Intelligence and Pattern Recognition*, 2016. https://doi.org/10.1109/ICAIPR.2016.7585225
25. G. Solana-Lavalle, J. Galán-Hernández, 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, *Biocybern. Biomed. Eng.*, **40** (2020), 505–516. https://doi.org/10.1016/j.bbe.2020.01.003
26. M. Nilashi, H. Ahmadi, A. Sheikhtaheri, R. Naemi, R. Naemi, R. Alotaibi, et al., Remote tracking of Parkinson's disease progression using ensembles of Deep Belief Network and Self-Organizing Map, *Expert Syst. Appl.*, **159** (2020). https://doi.org/10.1016/j.eswa.2020.113562
27. N. Fayyazifar, N. Samadiani, Parkinson's disease detection using ensemble techniques and genetic algorithm, in *IEEE Artificial intelligence and signal processing conference*, (2017), 162–165. https://doi.org/10.1109/AISP.2017.8324074
28. H. Kaur, A. Malhi, H. S. Pannu, Machine learning ensemble for neurological disorders, *Neural Comput. Appl.*, **32** (2020), 12697–12714. https://doi.org/10.1007/s00521-020-04720-1

29. S. Aich, K. Younga, K. Hui, A. Al-Absi, M. Sain, A nonlinear decision tree based classification approach to predict the Parkinson's disease using different feature sets of voice data, in *International Conference on Advanced Communication Technology*, (2018), 638–642. https://doi.org/10.23919/ICACT.2018.8323864
30. A. K. Dutta, N. M. A. Zakari, Y. Albagory, A. R. Wahab Sait, Colliding bodies optimization with machine learning based Parkinson's disease diagnosis, *Comput. Syst. Sci. Eng.*, **44** (2023), 2195–2207. https://doi.org/10.32604/csse.2023.026461
31. G. Prema Arokia Mary, N. Suganthi, Detection of Parkinson's disease with multiple feature extraction models and darknet CNN classification, *Comput. Syst. Sci. Eng.*, **43** (2022), 333–345. https://doi.org/10.32604/csse.2022.021164
32. R. Prashanth, S. Dutta Roy, P. K. Mandal, S. Ghosh, High-accuracy detection of early Parkinson's disease through multimodal features and machine learning, *Int. J. Med. Inf.*, **90** (2016), 13–21. https://doi.org/10.1016/j.ijmedinf.2016.03.001
33. F. Saeed, M. Al-Sarem, M. Al-Mohaimeed, A. Emara, W. Boulila, M. Alasli, et al., Enhancing Parkinson's disease prediction using machine learning and feature selection methods, *Comput. Mater. Continua*, **71** (2022). https://doi.org/10.32604/cmc.2022.023124
34. P. Magesh, R. Myloth, R. Tom, An explainable machine learning model for early detection of Parkinson's disease using LIME on DaTSCAN imagery, *Comput. Biol. Med.*, **126** (2020). https://doi.org/10.1016/j.compbiomed.2020.104041
35. M. Hires, M. Gazda, P. Drotar, N. Pah, M. Motin, D. Kumar, Convolutional neural network ensemble for Parkinson's disease detection from voice recordings, *Comput. Biol. Med.*, **141** (2022). https://doi.org/10.1016/j.compbiomed.2021.105021
36. M. A. Schulz, B. Yeo, J. T. Vogelstein, J. M-Miranada, J. N. Kather, K. Kording, et al., Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets, *Nat. Commun.*, **11** (2020). https://doi.org/10.1038/s41467-020-18037-z
37. K. Seddiki, P. Saudemont, F. Precioso, N. Ogrinc, M. Wisztorski, M. Salzet, et al., Cumulative learning enables convolutional neural network representations for small mass spectrometry data classification, *Nat. Commun.*, **11** (2020). https://doi.org/10.1038/s41467-020-19354-z
38. Y. Liu, Y. Li, X. Tan, P. Wang, Y. Zhang, Local discriminant preservation projection embedded ensemble learning based dimensionality reduction of speech data of Parkinson's disease, *Biomed. Signal Process. Control*, **63** (2021). https://doi.org/10.1016/j.bspc.2020.102165
39. Y. Qiu, H. Zheng, A. Devos, H. Selby, O. Gevaert, A meta-learning approach for genomic survival analysis, *Nat. Commun.*, **11** (2020). https://doi.org/10.1038/s41467-020-20167-3
40. M. R. Salmanpour, M. Shamsaei, A. Saberi, G. Hajianfar, H. Soltanian-Zadeh, A. Rahmim, Robust identification of Parkinson's disease subtypes using radiomics and hybrid machine learning, *Comput. Biol. Med.*, **129** (2021). https://doi.org/10.1016/j.compbiomed.2020.104142
41. A. Miladinovic, M. Ajcevic, P. Busan, J. Jarmolowska, G. Silveri, S. Mezzarobba, et al., Transfer learning improves MI BCI models classification accuracy in Parkinson's diseas e patients, in *European Signal Processing Conference*, 2021. https://doi.org/10.23919/Eusip co47968.2020.9287391
42. Q. Yu, Y. Ma, Y. Li, Enhancing speech recognition for Parkinson's disease patient using transfer learning technique, *J. Shanghai Jiaotong Univ.*, **27** (2022), 90–98. https://doi.org/10.1007/s12204-021-2376-3

43. H. Li, G. Wen, Sample awareness-based personalized facial expression recognition, *Appl. Intell.*, **49** (2019), 2956–2969. https://doi.org/10.1007/s10489-019-01427-2
44. Y. Gao, Y. Cui, Deep transfer learning for reducing health care disparities arising from biomedical data inequality, *Nat. Commun.*, **11** (2020). https://doi.org/10.1038/s41467-020-18918-3
45. Md. S. R. Sajal, Md. T. Ehsan, R. Vaidyanathan, S. Wang, T. Aziz, K. Mamun, Tele-monitoring Parkinson's disease using machine learning by combining tremor and voice analysis, *Brain. Inf.*, **7** (2020). https://doi.org/10.1186/s40708-020-00113-1
46. L. Zahid, M. Maqsood, M. Y. Durrani, M. Bakhtyar, J. Baber, H. Jamal, et al., A spectrogram-based deep feature assisted computer-aided diagnostic system for Parkinson's disease, *IEEE Access*, **8** (2020). https://doi.org/10.1109/ACCESS.2020.2974008
47. Y. Li, Y. Yang, S. Zhou, J. Qiao, B. Long, Deep transfer learning for search and recommendation, in *Companion Proceedings of the Web Conference*, (2020), 313–314. https://doi.org/10.1145/3366424.3383115
48. B. E. Sakar, G. Serbes, C. Okan Sakar, Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease, *PLoS. ONE*, **12** (2017), 1–18. https://doi.org/10.1371/journal.pone.0182428
49. K. Mamun, M. Alhussein, K. Sailunaz, M. Islam, Cloud based framework for Parkinson's disease diagnosis and monitoring system for remote healthcare applications, *Future Gener. Comput. Syst.*, **66** (2017), 36–47. https://doi.org/10.1016/j.future.2015.11.010
50. C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. 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, *Appl. Soft Comput.*, **74** (2019), 255–263. https://doi.org/10.1016/j.asoc.2018.10.022
51. M. Little, P. McSharry, E. Hunter, J. Spielman, L. Ramig, Suitability of dysphonia measurements for telemonitoring of Parkinson's disease, *Nat. Prec.*, **2008** (2008), 1015–1022. https://doi.org/10.1038/npre.2008.2298.1
52. T. Biloborodova, I. Skarga-Bandurova, I. Skarha-Bandurov, Knowledge and data acquisition in mobile system for monitoring Parkinson's disease, in *Information and Knowledge in Internet of Things*, Springer, (2022), 99–119. https://doi.org/10.1007/978-3-030-75123-4_5
53. L. Berus, S. Klancnik, M. Brezocnik, M. Ficko, Classifying Parkinson's disease based on acoustic measures using artificial neural networks, *Sensors*, **19** (2019), 1424–8220. https://doi.org/10.3390/s19010016

AIMS Press