








Classification of Handwritten Drawings of People with Parkinson's Disease by Using Histograms of Oriented Gradients and the Random Forest Classifier

João Paulo Folador¹ , Adrian Rosebrock² ,
Adriano Alves Pereira¹ , Marcus Fraga Vieira³ ,
and Adriano de Oliveira Andrade¹ 

¹ Centre for Innovation and Technology Assessment in Health, Postgraduate Program in Electrical and Biomedical Engineering, Faculty of Electrical Engineering, Federal University of Uberlândia, Uberlândia, Brazil
jpfolador@gmail.com, a.alves.pereira@uol.com.br,
adriano@ufu.br

² PyImageSearch, University of Maryland Baltimore County (UMBC),
Baltimore, USA
adrian@pyimagesearch.com

³ Bioengineering and Biomechanics Laboratory, Federal University of Goiás,
Goiânia, Brazil
marcus@ufg.br

Abstract. Parkinson's disease (PD) is a neurological disorder that is progressive and causes losses of dopaminergic neurons from the substantia nigra, a region in the human brain. The decrease of dopamine in this area elucidates the presence of motor symptoms, such as tremors, bradykinesia, rigidity, gait impairment, and non-motor symptoms, e.g., depression, loss of cognitive functions, sleep problems, and nerve pain. Among the motor symptoms, tremors can have the most impact on the social activities of people with PD. Furthermore, there is difficulty in diagnosing the underlying disorder that causes tremors. Thus, the study and development of methods to assess tremors and their severity is of paramount relevance for clinical practice. A typical clinical tool to evaluate tremor severity is the analysis of hand drawing shapes (e.g., spirals, circles, meanders, waves). The evaluation of these drawings is dependent on the experience of professionals, yielding a high variability of results. Aiming to contribute to the objective evaluation of hand drawing shapes of people with PD, this research proposes the application of the Random Forest Classifier to classify Histograms of Oriented Gradients (HOG) estimated from sinusoidal patterns collected from healthy individuals ($n = 12$) and from people with PD ($n = 15$). The highest accuracy, sensitivity and specificity classification success rates were of 83%, 85% and 81%, respectively. These results can be relevant for the early detection of pathological tremors, the follow-up of medical treatments and the diagnosis of parkinsonian conditions.

Keywords: Parkinson's disease · Tremor · Handwritten drawing classification · Random Forest Classifier · Histograms of oriented gradients

1 Introduction

Parkinson's disease (PD) is a neurological disorder that is progressive and causes losses of dopaminergic neurons from the substantia nigra, a region in the human brain. The decrease of dopamine in this area implies the worsening of motor symptoms such as tremors, bradykinesia, rigidity, gait impairment, and non-motor symptoms such as depression, loss of cognitive functions, sleep problems and nerve pain [1].

PD affects 1% of the world's population aged 60 years and over, and despite scientific advancement, the disease remains incurable. The diagnosis of PD is complex, with a seasoned specialist being necessary to make it [1, 2].

Tremors are a common symptom in PD and it can be classified into many types: resting tremor, postural tremor, kinetic, essential, cerebellar, and others. Each type manifests in different situations and frequency ranges [3].

Despite the existence of various clinical scales to assess motor symptoms in PD (e.g., Unified Parkinson's disease Rating Scale - UPDRS [2], Tremor Rating Scale - wTRS, and the Essential Tremor Rating Assessment Scale - TETRAS [3]), the understanding and quantification of tremors is important for the correct diagnosis of PD and for monitoring its progress [3, 4]. An alternative way to assess tremors is by using a scale of severity based on handwritten drawings. However, the scoring of these drawings is complex and dependent on the experience of the examiner.

Several methods have been proposed to automate the assessment of tremors in PD. For instance, Bravo et al. [5] analyzed postural tremor, action tremor and rest tremor from the hand index finger using a triaxial accelerometer to acquire the data. The data was analyzed via spectral power density (PSD). The study showed that the tremors decreased considerably with the use of medication, but they did not disappear completely.

Zhang et al. [6] employed principal component analysis (PCA) to discriminate the main features from Magnetic Resonance Imaging (MRI) data and Support Vector Machine (SVM) to classify the essential and PD tremors. The best classification success rate reported in the study was 93.75%.

Prince and de Vos [7] collected data from healthy individuals and people with PD, performing the task of hitting the index and middle fingers on a smartphone. Subsequently, they compared the data classification success rate between traditional algorithms and Deep Learning (DL), which outperformed traditional techniques.

According to Pereira et al. [8], another technique based on DL is Convolutional Neural Network (CNN), and it was used for the automatic discrimination between people with PD from healthy individuals. The data were acquired by using a pen with several attached sensors. The participants of the study drew spirals and meanders on a sheet of paper. The data from the sensors were converted into images for data classification and the CNN reached good accuracy.

Unlike the work done by Pereira et al. [8] that used the time series from different pens' sensors and built an image of these signals, this work proposes to classify images of handwritten drawings collected from healthy individuals and people with PD. The identification and discrimination of motor symptoms in PD is a fundamental step in the diagnosis and follow-up of the disorder

The remainder of this paper is organized as follows. Section 2 describes the experimental environment, the information about the participants of the study, and the feature extraction and classification methods. Section 3 shows the obtained results, and in Sect. 4 the discussion and conclusions are presented.

2 Materials and Methods

2.1 Computational Environment

The experiments were carried out in a machine with Intel Core i7 2.40 GHz, dual DDR3 with 8 GB RAM, 256 SSD of hard driver, and a 2 GB video card NVIDIA GeForce GT 650. The machine was configured with Microsoft Windows 7 Pro 64 bits, Python 3.6.5, the Scientific Python Development Environment (Spyder 3.3.2), and Keras that is a high-level API for building and training machine learning models.

2.2 Data Collection

Data was collected from 12 (twelve) healthy individuals and 15 (fifteen) individuals with Parkinson's disease. Table 1 shows information from the two groups. The Federal University of Uberlândia's Research Ethics Committee approved the research under the number 07075413.6.0000.5152.

Table 1. Characterization of the studied groups.

Group	Total	Sex (F/M)	Age (years)
Health	12	8/4	60.08 \pm 6.13
PD	15	7/8	65.33 \pm 9.17

The method based on severity scales was used to collect the data. In this method, the participants have to draw geometric shapes like spirals, sine waves, circles, or another different shape (e.g., Fig. 1).

2.3 Experimental Task

The participants involved in this research had to draw a specific image pattern similar to a sine wave. First, the person made the drawing following a printed pattern. A standard black pencil was used. After the participant learned how to draw the pattern, a new drawing was made, as illustrated in Fig. 1.

Each participant drew between three and four samples of sine waves. These drawings were digitalized, cleaned (the arrows were removed) and rescaled to a width of 512 pixels and automatic height (Gimp image manipulation software was used to preprocess the images).

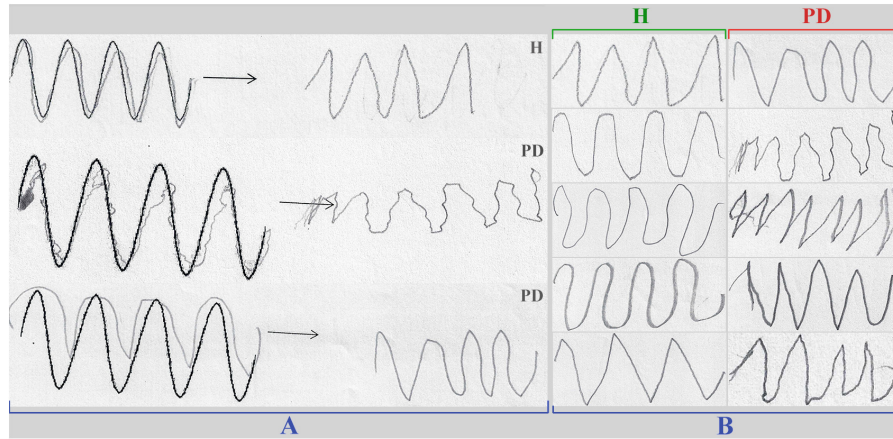


Fig. 1. (A) Samples of handwritten drawings collected from people with Parkinson's disease (PD) and healthy individuals (H); (B) Pre-processed images for each group (H and PD).

Figure 1(A) shows a sample of raw drawings made by a healthy person (H) and two distinct people with PD marked with (PD). Figure 1(B) shows drawing samples from each group.

In the study, 51 images were collected from each group, i.e., healthy individuals and people with PD. A total of 102 images were available.

2.4 Machine Learning

Machine learning is a subarea of artificial intelligence based on the idea that systems can learn from data, classify and identify patterns, and make decisions automatically. In this paper, we used some of these techniques to solve the problem of recognizing and classifying handwritten drawings between two different classes: drawings of healthy and Parkinson's disease subjects.

2.4.1 HOG Descriptor

The first step for image classification was to apply a method named histograms of oriented gradients (HOG). The HOG descriptor is commonly used to object detection. HOG allows the image to be described by the distribution of intensity gradients or edge directions. Figure 2 illustrates the wave detected with the intensity gradients and orientation [9, 10].

In Fig. 2 it is possible to notice that HOG divides the image into small areas named cells, which are of a predefined size, Fig. 2(B) in blue; the method estimates the histogram of the gradient orientations of each cell as shown in Fig. 2(C). Following this, normalization of the histograms in each cell is performed by comparing each block to the block of neighboring cells. Finally, a one-dimensional feature vector from the information in each cell is obtained [9–12]. The method scans and processes the entire image using the block to create the HOG that is presented in Fig. 2(D) as an output.

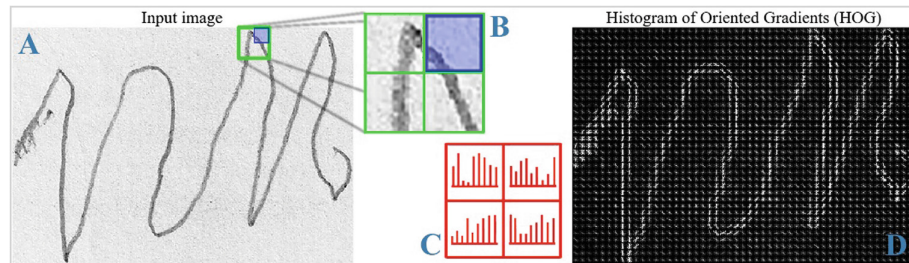


Fig. 2. Raw input image (A), the block of 2×2 cells (B), gradient histogram of each cell (C), and HOG feature extracted from the image input (D).

In this work, the input image was resized to 200 by 200 pixels (width and height). HOG was defined with 10 (ten) pixels per cell, blocks with 4 (four) cells (2×2 matrix), and the number of orientations was 9 (nine), meaning that nine bins were defined in the histogram with orientation between 0° and 180° degrees for each cell.

2.4.2 Random Forest Classifier

After HOG estimation, the data is ready to be classified by a Random Forest Classifier (RFC). RFC is a type of supervised machine learning algorithm based on ensemble learning, a method that makes it possible to join different types of algorithms or the same algorithm to set a more powerful prediction model. The random forest algorithm combines multiple decision-tree algorithms [13].

A decision tree (DT) is a tree in which a node represents a feature, each branch represents a decision and each leaf yields a result that can be a categorical or a continuous value [14, 15]. In addition, DT is a non-parametric supervised learning method commonly used for classification and regression [15, 16].

A Random Forest is a meta-estimator that fits multiple decision tree classifiers into manifold subsamples of the dataset and uses the mean to improve predictive accuracy and control overfitting. In general, an RFC takes N objects from the database, builds a decision tree with this data, and every tree in the forest predicts the category of the objects belonging to it. Finally, the new object is assigned to the category that wins the majority vote [13, 17, 18].

In this study, the Random Forest Classifier was structured with 100 and 200 decision trees and analyzed the model differences. The dataset was split into 70% of data for training and 30% for testing the model. Furthermore, the model was executed 10 times, 50 times and 100 times. Following that, an average of metrics was estimated and it was analyzed whether the model is able to classify the handwritten drawings of people with PD and healthy individuals.

2.4.3 Accuracy, Sensitivity and Specificity

These metrics are commonly used to describe if a test is good enough and reliable. The accuracy, sensitivity and specificity are the most used statistics to describe a diagnostic test [19]. Accuracy demonstrates the proportion of correct prediction of a given condition. Sensitivity evaluates how good the test is at detecting a positive disease. On the

other hand, the specificity shows us if a healthy subject has been correctly classified as without disease [16, 19].

The accuracy value is obtained by the number of correct assessments divided by the number of all assessments. Sensibility is calculated by the number of true positive assessments divided by the number of all positive assessments. Finally, the specificity is acquired through the number of true negative assessments divided by the number of all negative assessments [19].

3 Results

Table 2 describes the RFC results for each test. One test was arranged with 100 trees and it was executed in batches of 10, 50 and 100 times. For each one of these “total of runs”, the average of the classification obtained was made and the lowest and highest classification rates were estimated. For these tests the highest accuracy was 0.83 (83% of success) and the average was 70%. Sensitivity reached the best value of 83% and an average of 69%. The highest specificity was 85% and its mean was 70%.

Table 2. Classification results for distinct configurations of the Random Forest Classifier.

Set of trees	Total of runs	Accuracy			Sensitivity			Specificity		
		Min	Max	Average	Min	Max	Average	Min	Max	Average
100	10	0.63	0.76	0.70	0.64	0.83	0.71	0.61	0.79	0.70
	50	0.60	0.80	0.69	0.60	0.83	0.69	0.59	0.82	0.70
	100	0.57	0.83	0.70	0.57	0.82	0.69	0.56	0.85	0.70
200	10	0.67	0.77	0.72	0.67	0.75	0.70	0.67	0.79	0.74
	50	0.60	0.80	0.70	0.59	0.80	0.70	0.61	0.80	0.71
	100	0.60	0.80	0.70	0.61	0.80	0.70	0.58	0.80	0.72

The second configuration of the RFC was 200 trees. Table 2 shows the results. The highest value of accuracy was 80% and the average 71%. The highest sensitivity was 80% and the average was 70%. The highest specificity was 80% and the average was 72%.

The results are shown in Fig. 3 in confusion matrix (CM) format. It shows us the relation of true and false positives about the presence of tremor (T) in PD sufferers and the true and false negatives in healthy (H) subjects. The diagonal cells correspond to observations that are correctly classified, and the off-diagonal cells correspond to incorrectly classified observations. At the bottom right of the CM is the cell with the overall accuracy.

Figure 4 shows boxplots for each metric. Figure 4(A) shows a comparison between the accuracies obtained for three different batches for the RFC with 100 trees. A similar procedure was executed for sensitivity and specificity. Figure 4(B) presents results for the RFC with 200 trees.

Predicted		A1			A2			A3			
		T	H	Total	T	H	Total	T	H	Total	
		T	34.00%	16.00%	68.00%	34.93%	15.07%	69.87%	35.37%	14.63%	70.73%
		H	13.67%	36.33%	72.67%	15.60%	34.40%	68.80%	15.57%	34.43%	68.87%
Total		71.33%	69.43%	70.33%	69.13%	69.54%	69.33%	69.44%	70.18%	69.80%	
		28.67%	30.57%	29.67%	30.87%	30.46%	30.67%	30.56%	29.82%	30.20%	
Target Class											
Predicted		B1			B2			B3			
		T	H	Total	T	H	Total	T	H	Total	
		T	38.00%	12.00%	76.00%	36.33%	13.67%	72.67%	36.60%	13.40%	73.20%
		H	16.33%	33.67%	67.33%	15.73%	34.27%	68.53%	15.77%	34.23%	68.47%
Total		69.94%	73.72%	71.67%	69.78%	71.49%	70.60%	69.89%	71.87%	70.83%	
		30.06%	26.28%	28.33%	30.22%	28.51%	29.40%	30.11%	28.13%	29.17%	
Target Class											

Fig. 3. Confusion matrix test average. A1, A2, and A3 are from the RFC tests with 100 trees with 10, 50, and 100 runs, respectively. In the same way, RFC tests with 200 trees are represented by B1, B2, and B3, respectively in 10, 50, and 100 runs.

The data distribution in Fig. 4(A) is around 65–75% for all metrics. In Fig. 4(B) the accuracy and sensitivity behave the same way in (A), and specificity spread above 75%.

4 Discussion and Conclusion

In this work, the drawings collected from healthy individuals and people with PD were classified by RFC. The proposed method employed pencil drawings digitized from ordinary sheet of paper, making it very simple to be applied in the context of scarce financial resources. A major advantage of RFC is the low computational cost when compared to Deep learning.

Despite the small number of images in the available data set (51 per class), the obtained results were satisfactory and accurate by discriminating drawings of healthy people from those with PD (Fig. 3).

In the study, the HOG parameters were tested in default values (10×10 pixels per cell, 2×2 cells per block and 9 bins in the histogram with $0\text{--}180^\circ$ orientation) focus on good performance showed by Dalal and Triggs [10] and the HOG result was passed to the classifier. In a future study, these parameters could be changed aiming to get the best ones to improve the model results.

The results shown in Table 2 and Fig. 4 suggest that there is a similarity regarding the number of trees used (100 and 200). The variability of all the metrics showed in

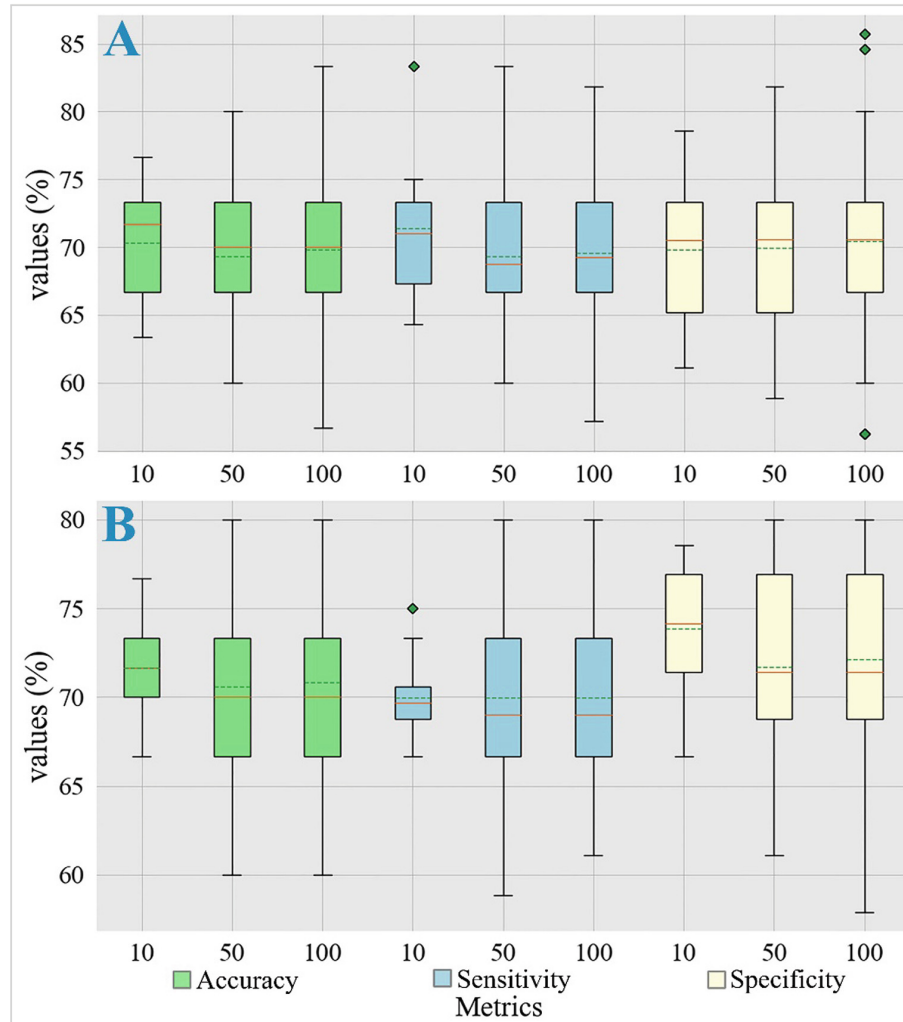


Fig. 4. Accuracy, sensitivity and specificity for each performance metric. (A) RFC with 100 trees and (B) RFC setup with 200 trees. The green dotted line is the average and the orange is the median.

Fig. 4 indicates that the diagnostic test is working correctly in discriminating between who has tremors and who does not.

This is the first reported study considering the application of HOG estimates in combination with the RFC applied to the automatic classification of data obtained from people with PD. This study is in the direction of related work [16] which analyzed data of people with dementia.

In the future, it will be necessary to obtain more image drawings and different shapes to increase the database [8]. In addition, it is relevant to test more parameters

and tune the proposed method as well as to implement other types of classifiers and compare them with RFC as proposed here.

Acknowledgment. The present work was carried out with the support of the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES – Program CAPES/DFATD-88887.159028/2017-00) and the Foundation for Research Support of the State of Minas Gerais (FAPEMIG-APQ-00942-17). A. O. Andrade, A. A. Pereira and M. F. Vieira are a fellow of CNPq, Brazil (304818/2018-6, 310911/2017-6, and 306205/2017-3).

Conflicts of Interest. The authors declare no conflict of interest.

References

1. de Lau, L.M., Breteler, M.M.: Epidemiology of Parkinson's disease. *Lancet Neurol.* **5**, 525–535 (2006). [https://doi.org/10.1016/S1474-4422\(06\)70471-9](https://doi.org/10.1016/S1474-4422(06)70471-9)
2. Cancela, J., Mascato, S.V., Gatsios, D., et al.: Monitoring of motor and non-motor symptoms of Parkinson's disease through a mHealth platform. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, pp. 663–666. IEEE (2016)
3. Andrade, A.O., Pereira, A.A., de Almeida, M.F.S., et al.: Human tremor: origins, detection and quantification. In: Andrade, A.O. (ed.) *Practical Applications in Biomedical Engineering*. InTech, Croatia (2013)
4. Chan, P.Y., Ripin, Z.M., Halim, S.A., et al.: An in-laboratory validity and reliability tested system for quantifying hand-arm tremor in motions. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**, 460–467 (2018). <https://doi.org/10.1109/TNSRE.2017.2782361>
5. Bravo, M., Bermeo, A., Huerta, M., et al.: A system for finger tremor quantification in patients with Parkinson's disease. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, pp. 3549–3552. IEEE (2017)
6. Zhang, L., Liu, C., Zhang, X., Tang, Y.Y.: Classification of Parkinson's disease and essential tremor based on structural MRI. In: Proceedings - 2016 7th International Conference on Cloud Computing and Big Data, CCBBD 2016, pp. 353–356. IEEE (2017)
7. Prince, J., de Vos, M.: A deep learning framework for the remote detection of Parkinson's disease using smart-phone sensor data. In: Proceedings of the Annual International Conference on IEEE Engineering in Medicine and Biology Society, pp. 3144–3147 (2018). <https://doi.org/10.1109/embs.2018.8512972>
8. Pereira, C.R., Weber, S.A.T., Hook, C., et al.: Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: Proceedings - 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI 2016, pp. 340–346. IEEE (2017)
9. Chowdhury, S.A., Kowsar, M.M.S., Deb, K.: Human detection utilizing adaptive background mixture models and improved histogram of oriented gradients. *ICT Express* **4**, 216–220 (2018). <https://doi.org/10.1016/j.ict.2017.11.016>
10. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), pp. 886–893. IEEE (2018)

11. Li, T., Li, W., Yang, Y., Zhang, W.: Classification of brain disease in magnetic resonance images using two-stage local feature fusion. *PLoS One* **12**, 1–19 (2017). <https://doi.org/10.1371/journal.pone.0171749>
12. Ren, H., Li, Z.-N.: object detection using edge histogram of oriented gradient
13. Breiman, L.: Random forests. *Mach. Learn.* **45**, 1–33 (2001)
14. Quinlan, J.R.: Induction of decision trees. *Mach. Learn.* **1**, 81–106 (1986). <https://doi.org/10.1007/BF00116251>
15. Safavian, S.R., Landgrebe, D.: A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **21**, 660–674 (1991). <https://doi.org/10.1109/21.97458>
16. Maroco, J., Silva, D., Rodrigues, A., et al.: Data mining methods in the prediction of Dementia: a real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res. Notes* **4**, 299 (2011). <https://doi.org/10.1186/1756-0500-4-299>
17. Bernard, S., Heutte, L., Adam, S.: Influence of hyperparameters on random forest accuracy. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 171–180. Springer (2009)
18. Ramani, R.G., Sivagami, G.: Parkinson disease classification using data mining algorithms (2011)
19. Zhu, W., Zeng, N., Wang, N., et al.: Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG Proc. Heal Care Life Sci. Balt. Maryl.* **19**, 67 (2010)