



# A Model-Free Computer-Assisted Handwriting Analysis Exploiting Optimal Topology ANNs on Biometric Signals in Parkinson's Disease Research

Vitoantonio Bevilacqua<sup>1</sup>(✉), Claudio Loconsole<sup>1</sup>, Antonio Brunetti<sup>1</sup>, Giacomo Donato Cascarano<sup>1</sup>, Antonio Lattarulo<sup>1</sup>, Giacomo Losavio<sup>2</sup>, and Eugenio Di Sciascio<sup>1</sup>

<sup>1</sup> Department of Electrical and Information Engineering (DEI),  
Polytechnic University of Bari, Bari, Italy  
vitoantonio.bevilacqua@poliba.it  
<sup>2</sup> Medica Sud s.r.l., Bari, Italy

**Abstract.** In this paper, we propose a novel model-free technique for differentiating both Parkinson's Disease (PD) patients from healthy subjects and mild PD patients from moderate ones by using a handwriting analysis tool. The tool is based on the analysis of biometric signals and the application of Artificial Neural Network (ANN)-based classifier. Experimental tests have been carried on with both healthy and PD subjects to identify the most representative features and to assess the accuracy and repeatability of classification performances achieved through optimal topology ANNs. Finally, the obtained results are reported and discussed to infer some important properties on classification approaches and the role of muscular activities on the handwriting analysis applied to neurodegenerative disease research.

**Keywords:** Handwriting analysis · Model-free · sEMG · Parkinson Disease ANN · MOGA

## 1 Introduction

The handwriting is a highly overlearned fine and complex manual skill involving an intricate blend of cognitive, sensory and perceptual-motor components [1]. For these reasons, the presence of abnormality in the handwriting process is a well-known and well-recognized manifestation of a wide variety of neuromotor diseases. There are two main difficulties related to the handwriting and affecting Parkinson's Disease (PD) patients: (i) the difficulty in controlling the amplitude of the movement, i.e., decreased letter size (micrographia) and failing in maintaining stroke width of the characters as writing progresses [2], and (ii) the irregular and bradykinetic movements, i.e., increased movement time, decreased velocities and accelerations, and irregular velocity and acceleration trends over time [3]. For these reasons, in literature, there are several works investigating the possibility of a differentiation between PD patients and healthy subjects by means of computer-aided handwriting analysis tools [3, 4].

In our previous work [5], we proposed a preliminary approach for differentiating PD patients from healthy subjects using a reduced set of features (4 features) by exploiting computer vision techniques applied on the scan of common paper sheets and surface ElectroMyoGraphy (sEMG) signal processing. We found dynamic features being more representative for the differentiation. In a subsequent work [6], instead, we used a graphic tablet and a sEMG bracelet during handwriting tasks to respectively extract biometric signals related to pen movements and to muscular activity. The resulting performance allowed us to assess both the comparison among different classification approaches and the differentiation between PD patients and healthy age-related subjects.

In this paper, we improved our previous research extending both the number and type of proposed features and the subject dataset. We have also investigated the most representative features to be used in handwriting research applied to PD and the differentiation between mild and moderate PD patients.

## 2 The Proposed Model-Free Technique

### 2.1 Handwriting Feature Extraction

The features related to handwriting were extracted from biometric signals acquired during the handwriting tasks. In particular, it is possible to group the proposed features into two categories - sEMG related and pen tip related features:

- *sEMG related features* – these features are related to the muscular activity of the subject and are extracted from the sEMG signals acquired at the subject’s forearm:
  - *Root Mean Square (RMS) features* extracted for each sEMG channel. RMS is computed as the square root of the mean of the sample squares.
  - *Zero Crossing (ZC) features*, an index related to the signal sign variation. To normalize the features among the subjects, its value is divided by the length of the signal.
- *Pen tip related features* - these features are extracted from the signals generated by a graphic tablet during the handwriting task:
  - *Cartesian and XY features* are referred to the pen tip position and are extracted starting from the XY axes position: *Cartesian and XY (i) velocity, (ii) acceleration, and (iii) jerk*. This lead to a total of nine signals.
  - *Pen tip pressure feature*, a scalar feature and corresponds to the pressure applied by the pen tip on the surface of the tablet.
  - *Azimuth and altitude feature*: the azimuth feature is the value of the angle between a reference direction (e.g., the Y axes of the tablet) and the pen direction projected on the horizontal plane. The altitude feature is the value of the angle between the pen direction and the horizontal plane.
  - *Pattern specific features* associated to a specific writing pattern (WP). For letter-based WPs, the features are mainly related to the writing size, whereas for the spiral-based WPs, the features are mainly related to the writing precision. For the features extracted from the letter-based WPs, the upper and the lower peaks of the Y coordinate of the pen tip position are computed and, then, used as input

data of a linear regressor. Finally, the angle  $\alpha$  between the  $R_{up}$  and  $R_{low}$  regression line and the coefficient of determination ( $R^2$ ) are computed and selected as features. For spiral WPs, instead, the feature extracted is an index representative of the variability of the strokes. For each point P of the X-Y pen tip position, the vector  $\vec{r}$  with respect to the spiral centroid point C having origin in P is computed. The angle  $\beta$  between  $\vec{r}$  and the direction vector  $\vec{d}$  tangent to the spiral in P is, then, calculated. The spiral precision index feature is the standard deviation of the  $\beta$  angles computed for each point P.

## 2.2 Feature Selection and Classification

To reduce the number of features to be classified and to infer which of them are the most representative of the subject's status, we used a classification decision tree technique based on Gini's diversity index. To classify the extracted features, we used Artificial Neural Network (ANN) based classifier. The optimal topology for an ANN classifier was found by exploiting a Multi-Objective Genetic Algorithm (MOGA) and by maximizing the average test accuracy on a certain number of training, validation and test iterations for each ANN topology using different permutations of the dataset [7].

The performance for both the MOGA algorithm and the ANN-based classification were evaluated in terms of accuracy, specificity and sensitivity.

## 3 Experiments and Results

### 3.1 Participants

32 participants (21 males, 11 females, age:  $71.4 \pm 8.3$  years old) took part in the experimental tests. In detail, the age-matched control group was composed of 11 healthy subjects (4 males, 7 females, age:  $70.2 \pm 10.2$  years old), whereas the PD group was composed of 21 subjects (17 males, 4 females, age:  $72.1 \pm 8.3$ ). According to the degree of the disease, the PD group was following divided into two subgroups: mild and moderate. The mild group was composed of 12 patients (9 males, 3 females, age:  $70.5 \pm 10.0$ ), whereas the moderate one was composed of 9 patients (8 males, 1 female, age:  $73.8 \pm 6.0$ ).

### 3.2 System Setup

The system setup for data acquisition is reported in Fig. 1. It includes two main sensors: (i) the Myo™ Gesture Control Armband allowing us to synchronously acquire 8 different sEMG sources at the forearm, and (ii) the WACOM Cintiq 13" HD, a graphics tablet providing visual feedback for acquiring pen tip planar coordinates and pressure, and the tilt of the pen with respect to the writing surface.



**Fig. 1.** Example of the system set-up used for data acquisition

### 3.3 Experimental Description

For the experiments, we used three writing patterns (WPs) leading to as many writing tasks, these are: a five-turn spiral drawn in anticlockwise direction (WP 1), a sequence of 8 Latin letter “l” with a size of 2.5 cm (WP 2) and with a size of 5 cm (WP 3). Since the last two WPs were size-constrained, a visual marker was provided as reference. In the experiment, we asked each subject to perform the three writing tasks four times each for a total of twelve tasks: first for familiarization purposes, whereas the other three were acquired and stored for the subsequent feature extraction and processing. The subject was asked to rest between two subsequent handwriting tasks for at least three seconds. The beginning of the task signal acquisition was triggered by a positive pen pressure applied on the graphic tablet. The processing of the acquired raw signals led to the extraction of 41 features for WP 1 and to 43 features for WP 2 and 3.

### 3.4 Experimental Data Processing Description

We conducted the experiments under two main objectives: (i) the separation of the PD patients from healthy ones, and (ii) the classification of mild and moderate Parkinson subjects. The following features extracted during the experiments, conducted according to Sect. 3.3, were grouped in three datasets: (i) dataset A with 41 features, (ii) dataset B with 43 features and (iii) dataset C with 43 features extracted from WP 1, WP 2 and WP 3, respectively. Then, a feature selection algorithm was applied on the three datasets to select and reduce the number of the features. This led to the creation of six new different feature datasets: dataset with all features included in set A, B and C (Case 1, 2 and 3, respectively) and dataset with only the features resulting from the feature selection algorithm applied on dataset A, B and C (Case 4, 5 and 6, respectively).

### 3.5 Results and Discussion

Since we performed 250 iterations of the net training procedure for each case, the performance results have been reported in percentage (standard deviation in brackets).

*Objective 1 - Separating PD patients and healthy subjects:*

- for dataset A (WP 1–41 features – Case 4), the 6 selected features were: one RMS value, 3 ZC values, the mean cartesian velocity and acceleration on X axes;
- for dataset B (WP 2–43 features – Case 5), the 6 selected features were: the mean jerk on Y axes, 3 ZC values, the mean cartesian acceleration and velocity on X axes;

- for dataset C (WP 3–43 features – Case 6), the 7 selected features were: 2 RMS values, one ZC value, the mean cartesian velocity, the altitude STD, the azimuth RMS and the mean velocity on X axes.

The best accuracy value (96.85%) was achieved in case 6 (classification on the dataset composed of the selection of 7 features from the dataset of 43 features extracted from WP 3, i.e., the sequence of 8 Latin letter “I” with a size of 5 cm). In case 6, three out of seven features were related to sEMG signals (RMS and ZC), whereas the other features were related to pen tilt and velocity.

*Objective 2 - Separating mild and moderate PD patients:*

- for dataset A (WP 1–41 features – Case 4), the 6 selected features were: 2 RMS values, 2 ZC values, the mean pressure and the mean altitude;
- for dataset B (WP 2–43 features – Case 5), the 5 selected features were: 2 RMS values, 2 ZC values and the mean cartesian velocity;
- for dataset C (WP 3–43 features – Case 6), the 5 selected features were: 2 RMS values, one ZC value, the mean cartesian velocity on X axes and the mean pressure.

The best accuracy value (96.00%) was achieved in Case 4 (dataset A - 6 features selected over 41 features extracted from WP 1, i.e., the spiral WP). In Case 4, four out of six features were related to sEMG signals (RMS and ZC), whereas the other features were related to pen tilt and pressure.

The obtained classification accuracy for all four cases for both objectives, instead, are reported in Table 1. As it can be observed, the obtained accuracy values  $x$  for both objectives are high ( $86 < x < 97$ ) and present a limited standard deviation  $d$  ( $d < 0.09$ ), thus demonstrating the repeatability of the classification performances and the stability of the optimal topology ANN architectures. It is worth to observe also that the highest values of resulting accuracy have been obtained for both objectives for the classification of the selected features. The obtained results allow us to confirm the relevance of the sEMG signals not only in differentiating PD patients and healthy patients, but also, and especially, in differentiating mild and moderate PD patients. Furthermore, the results confirm the choice of acquiring signals related also to pen tilt, pressure and velocity.

**Table 1.** Accuracy and standard deviation values obtained for each case.

	Case	Objective 1	Objective 2
All features	1	90.76% (0.0764)	94.34% (0.0626)
	2	92.98% (0.0523)	87.26% (0.0850)
	3	95.95% (0.0479)	91.86% (0.0830)
Selected features	4	93.78% (0.0566)	96.00% (0.0658)
	5	91.58% (0.0526)	86.71% (0.0837)
	6	96.85% (0.0405)	91.66% (0.0858)

## 4 Conclusion

In this work, we proposed a model-free technique for computer-assisted handwriting analysis. The technique is based on the extraction of different features from biometric signals acquired during the handwriting task and on their classification by means of optimal topology ANNs whose characteristics result from a MOGA processing.

The proposed technique allowed us to tackle two main research objectives: the differentiation between healthy subjects and PD patients (objective 1) and between mild and moderate PD patients (objective 2) both with a high classification accuracy (over 90%). Furthermore, we demonstrated that a limited number of representative feature, selected by means of a classification decision tree technique based on Gini's diversity index, allowed us to obtain a more performing classification for both the objectives of the study (up to 96.85% and 96.00% for objective 1 and 2, respectively).

**Acknowledgments.** This work has been partially funded from the FutureInResearch program of the Regione Puglia - project n. JTFWZV0 ABIOSAN - Advanced BIometric analysis Against Neuromuscular disease.

## Bibliography

1. Carmeli, E., Patish, H., Coleman, R.: The aging hand. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* **58**, M146–M152 (2003)
2. Van Gemmert, A.W.A., Teulings, H.-L., Contreras-Vidal, J.L., Stelmach, G.E.: Parkinsons disease and the control of size and speed in handwriting. *Neuropsychologia* **37**, 685–694 (1999)
3. Drotar, P., Mekyska, J., Smekal, Z., Rektorova, I., Masarova, L., Faundez-Zanuy, M.: Prediction potential of different handwriting tasks for diagnosis of Parkinson's. In: 2013 E-Health and Bioengineering Conference (EHB), pp. 1–4 (2013)
4. Rosenblum, S., Samuel, M., Zlotnik, S., Erikh, I., Schlesinger, I.: Handwriting as an objective tool for Parkinson's disease diagnosis. *J. Neurol.* **260**, 2357–2361 (2013)
5. Loconsole, C., et al.: Computer vision and EMG-based handwriting analysis for classification in Parkinson's disease. In: Huang, D.-S., Jo, K.-H., Figueroa-Garcia, J.C. (eds.) ICIC 2017. LNCS, vol. 10362, pp. 493–503. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-63312-1\\_43](https://doi.org/10.1007/978-3-319-63312-1_43)
6. Loconsole, C., Cascarano, G.D., Brunetti, A., Trotta, G.F., Losavio, G., Bevilacqua, V., Di Sciascio, E.: A model-free technique based on computer vision and sEMG for classification in Parkinson's disease by using computer-assisted handwriting analysis. *Pattern Recognit. Lett.* (2018, in press). <https://doi.org/10.1016/j.patrec.2018.04.006>
7. Bevilacqua, V., Brunetti, A., Triggiani, M., Magaletti, D., Telegrafo, M., Moschetta, M.: An optimized feed-forward artificial neural network topology to support radiologists in breast lesions classification. In: Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion - GECCO 2016 Companion, pp. 1385–1392 (2016)