Benchmarking of Shallow Learning and Deep Learning Techniques with Transfer Learning for Neurodegenerative Disease Assessment Through Handwriting

Vincenzo Dentamaro ⓘ, Paolo Giglio ⓘ, Donato Impedovo(✉) ⓘ, and Giuseppe Pirlo ⓘ

University of Bari "Aldo Moro", Via Orabona 4, Bari, Italy
donato.impedovo@uniba.it

Abstract. Neurodegenerative diseases are incurable diseases where a timely diagnosis plays a key role. For this reason, various techniques of computer aided diagnosis (CAD) have been proposed. In particular handwriting is a well-established diagnosis technique. For this reason, an analysis of state-of-the-art technologies, compared to those which historically proved to be effective for diagnosis, remains of primary importance. In this paper a benchmark between shallow learning techniques and deep neural network techniques with transfer learning are provided: their performance is compared to that of classical methods in order to quantitatively estimate the possibility of performing advanced assessment of neurodegenerative disease through both offline and online handwriting. Moreover, a further analysis of their performance on the subset of a new dataset, which makes use of standardized handwriting tasks, is provided to determine the impact of the various benchmarked techniques and draw new research directions.

Keywords: Shallow learning · Deep learning · Neurodegenerative disease · Handwriting · Deep neural networks · Benchmark

1 Introduction

Several non-invasive techniques have been developed in order to assess the presence of neuro-degenerative diseases, which is characterized by a gradual decline of cognitive, functional and behavioral areas of the brain [1, 2]. Among them, behavioral biometrics, such as speech [4], have proven to be promising in terms of accuracy in binary classification (healthy/unhealthy) for neurodegenerative diseases assessment. Handwriting behavioral biometric particularly stands out for its strict relation with the level of severity of a vast class of neurodegenerative diseases, therefore its features' changes are considered an important biomarker: [1, 2] indeed handwriting involves kinesthetic, cognitive and perceptual-motor tasks [4], resulting in a very complex activity whose performance is taken into account for the evaluation of several diseases such as PD and AD [3, 5–7].

This work proposes a benchmark of traditional shallow learning techniques with deep learning techniques for neurodegenerative disease assessment though handwriting.

This work consists of handwriting acquisitions performed online via tablet: variables like x, y coordinates as well as azimuth, pressure, altitude, in air movements and timestamps of each acquisition are collected. For the specific purpose of the study, only the final handwritten trace, i.e. the whole set of x,y coordinates and the azimuth, is used as the training data set. The handwriting procedure consists of 8 different tasks which will be show in detail in Sect. 5. The paper is organized as follows. Section 2 sketches state of art review for neurodegenerative disease assessment through handwriting, Sect. 3 illustrates the use of shallow learning technique on on-line handwriting recognition by means of velocity based-features and kinematic-based features. In Sect. 4 both offline and online deep learning techniques are presented. Section 5 shows dataset description and results. Reasoning of results is provided in Sect. 6. Finally, Sect. 7 sketches conclusions and future remarks.

2   State of the Art Review

The aim of this work is to provide insights about the best features and techniques to adopt into a computer aided diagnosis system for supporting early diagnosis of neurodegenerative disease. It is important to not only predict the disease, but also to monitor the progression of it during time. [1, 2]. The scientific community focused the research towards predictive models that can accurately detect subtle changes in writing behavior. These techniques will be used to help neurologists, and psychologists to assess diseases as an auxiliary tool in addition to the battery of cognitive tests provided in literature [1–9].

The acquisition tool, at time of writing, is a digital tablet with a pen. This device captures spatial and temporal data and save it inside a storage memory. After data is captures, as often happens in shallow learning scenario, features are extracted. Usually, patients are asked to perform several tasks [1].

Even though important results were achieved by the community, there is not homogeneity in tasks provided by the datasets developed. That is because scientists collected databases of handwriting tasks themself resulting in datasets with different kind of tasks, usually not connected among them and merged together, which provided controversary results. To overcome this problem, authors in [34] have developed a specific acquisition protocol. This protocol includes a digitizer version of standard tests used, accepted, tested in the neurological community used as the ground truth for evaluation. The dataset used in this work is a subset of this big dataset which is currently under development. This dataset contains well-established handwriting tasks to perform kinematic analysis and handwriting experimental tasks useful for extracting novel types of features to be investigated by researchers. Literature review on handwriting recognition for neurodegenerative disease assessment can be subdivided in two main groups: online handwriting and offline. In the online handwriting, the features computed in all the tasks are then concatenated into a high dimensional vector and then used for classification. [9] Various authors used several kinds of classifiers ranging from SVM, KNN, ensemble learning with Random Forests, neural networks and so on. [1–9]. It has also been analyzed the

use of an ensemble of classifiers each one built onto one single feature space of each task [1, 2].

For the online handwriting recognition for neurodegenerative disease assessment, some of the authors of this work in [1] used several features like position, button status, pressure, azimuth, altitude, displacement, velocity and acceleration over 5 different datasets namely: PaHaw [9], NewHandPH [29], ParkinsonHW [30], ISUNIBA [31], EMOTHAW[32] achieving accuracies that range from 79.4% to 93.3% depending on the dataset and tasks.

For offline handwriting recognition, authors in [33] used "enhanced" static images of handwriting generated by exploiting simultaneously the static and dynamic properties of handwriting by drawing the points of the samples and adding pen-ups for the same purpose. Authors used a Convolutional Neural Network to provide feature embedding and then a set of classifier is used in a majority voting fashion. Authors used transfer learning for coping with limited amount of training data. Their accuracies on the various tasks ranged from 50% to 65% showing some limits of this technique.

In [35] authors explored an alternative model that used one single bi-directional LSTM layers on handwriting recognition tasks, achieving better or equivalent results than stacking more LSTM layers, which decreases the complexity and allows a faster network training. In [36] authors investigated the use of bidirectional LSTM with attention mechanism for offline and online handwriting recognition achieving important results on the RIMES handwriting recognition task. The bidirectional LSTM architecture developed in this work was partially inspired by the work in [36]. Some of the authors of this work used also computer vision for assessing neurodegenerative disease through gait [37] and sit to stand tasks [38].

3 Shallow Learning for Online Handwriting Neurodegenerative Disease Assessment

The term shallow learning identifies all techniques that do not belong to deep learning. In the case of online handwriting recognition, where online stands for capturing time-series of movements of the pen on a digital support, shallow learning is equivalent to perform feature extraction and classification with various machine learning algorithms. To this extent, standard velocity-based features and kinematic-based features are extracted and tested with random forest classification algorithm. The set of extracted features is shown in Table 1. All the features extracted were standardized. Moreover, Random Forest [17] ensemble learning algorithm with features ordered by relevance was adopted to make a selection of the most important criteria [18]. The random forest pre-pruning parameter was of maximum tree depth of 10 and 50 trees, in order to prevent overfitting and to balance accuracies. Random Forest algorithm [17] was also used for classification purposes: its maximum depth adopted was of 10 and the number of trees was estimated dynamically with the inspection of the validation curve. The reported accuracies are based on a 10-fold cross validation, i.e. the entire procedure was repeated 10 times, where each fold was used as a test set.

Table 1. Features used in shallow learning.

Feature	Description
Position	Position in terms of s(x,y)
Button status	Movement in the air: $b(t) = 0$ Movement on the pad: $b(t) = 1$
Pressure	Pressure of the pen on the pad
Azimuth	Angle between the pen and the vertical plane on the pad
Altitude	Angle between the pen and the pad plane
Displacement	$d_i = \begin{cases} \sqrt[2]{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}, & 1 \le i \le n-1 \\ d_n - d_{n-1}, & i = n \end{cases}$
Velocity	$v_i = \begin{cases} \frac{d_i}{t_{i+1} - t_i}, & 1 \le i \le n-1 \\ v_n - v_{n-1}, & i = n \end{cases}$
Acceleration	$a_i = \begin{cases} \frac{v_i}{t_{i+1} - t_i}, & 1 \le i \le n-1 \\ a_n - a_{n-1}, & i = n \end{cases}$
Jerk	$j_i = \begin{cases} \frac{a_i}{t_{i+1} - t_i}, & 1 \le i \le n-1 \\ j_n - j_{n-1}, & i = n \end{cases}$
x/y displacement	Displacement in the x/y direction
x/y velocity	Velocity in the x/y direction
x/y acceleration	Acceleration in the x/y direction
x/y jerk	Jerk in the horizontal/vertical direction
NCV	Number of Changes in Velocity, NCV has been also normalized to writing duration
NCA	Number of Changes in Acceleration, NCA has been also normalized to writing duration

3.1 Velocity-Based Features

The choice of certain velocity-based features is dictated by motor deficits particularly present in neurodegenerative diseases. Motor deficits like bradykinesia (which is characterized by slowness of movements), micrographia (time related reduction of the size of writing), akinesia (characterized by impairment of voluntary movements), tremor and muscular rigidity [2], are particularly evident when patient is asked to perform certain tasks. These tasks are often characterized by drawing stars, spirals, writing names and copying tasks [3, 8–10, 12]. In order to model other symptoms such as tremor and jerk, the patient is often asked to draw meanders, horizontal lines, straight (both forward and backward) slanted lines, circles and few predefined sentences as shown in [11–13].

Table 1 shows the features extracted for the shallow learning classification. It is important to state that every feature is a time-series, thus statistical functions such a mean, median, standard deviation 1st and 99th percentile are used to synthetize each feature in Table 1.

3.2 Kinematic-Based Features

For modelling online handwriting and extracting important movements patterns, authors in [14] have used the Maxwell-Boltzmann distribution. This distribution is used to extract parameters that are then used to model the velocity profile. Its formulation is shown in formula (1).

$$mb_j = v_j^2 e^{-v_j^2} \tag{1}$$

V_j is the velocity at j-th position. Another kinematic feature used for describing the handwriting pattern of velocity and acceleration profile is the Discrete Fourier Transform as shown in [15]. Its formulation, shown in (2), is composed by the computation of the DFT and the computation of the Inverse DFT, which contains the spectrum of harmonics having the magnitude inversely proportional to the frequency [16]. Thanks to the logarithm present in the formulation, components with small variations tends to converge toward 0, instead repeated peaks at higher frequencies are typical of periodic patterns described by tremor and jerks.

$$rcep = IDFT\left\{log\left[\left|DFT\left(v_j\right)\right|\right]\right\} \tag{2}$$

Again, v_j is the velocity at j-th position.

4 Deep Learning for Offline and Online Handwriting Neurodegenerative Disease Assessment

Deep Learning techniques have been developed for various tasks such as image recognition through convolutional neural networks, but also time series analysis using recurrent neural networks with stacked layers such as LSTM and bi directional LSTM. The motivation behind this work is to benchmark deep learning architectures trained by using deep transfer learning on images generated by x,y coordinates drawing, from now on referred as offline handwriting, with respect to online handwriting models trained on time series of x,y coordinates for the RNN and shallow learning as reported in Sect. 3.

4.1 CNN Based Networks for Offline Recognition

For the offline handwriting recognition, 224 by 224 pixels images are generated by plotting x, y coordinates of each task and saving the generated image.

Because of the limited amount of training data, it has been decided to use deep transfer learning [20]. Deep transfer learning is useful when not much training data is available, as in this case. The idea is to use a deep neural network architecture and

weights trained on a big dataset and fine tune, only final layers on our dataset by freezing the former layers. This is useful, because initial layers usually generate high level representation of the underlying patterns, while the last layers are specialized in applying the proper classification. [20] All the used deep learning architectures are originally trained on Imagenet dataset [21] and then a 2D global average pooling layer has been added followed by one dense layer with 32 neurons and ReLU activation function, and finally the softmax layer for performing binary classification. New added layers are trained on the training set for 100 epochs and cross validated on the 33% of the training set.

All labels are one-hot encoded. The following architectures were chosen depending on the importance in the literature, disk size, number of parameters and the accuracy achieved on Imagenet dataset as reported on Keras website [22]. The chosen architectures are briefly reported are in Table 2.

Table 2. Deep learning architectures used

Architecture name	Size	Parameters	Top-5 accuracy on imagenet
NASNetLarge	343 MB	88,949,818	0.960
ResNET 50	98 MB	25,636,712	0.921
Inception V3	92 MB	23,851,784	0.937
Inception ResNet V2	215 MB	55,873,736	0.953

4.1.1 NASNetLarge

The architecture of NASNet Large deep neural network was not invented by a human being, but is the result of a process called Neural Architecture Search, where parameters of the network and its architecture is discovered as the output of an optimization process which uses reinforcement learning to learn and decide what is the best choice of layer type and hyperparameters given a specific dataset. In authors experiments [23], the algorithm searched for the best convolutional layer (or "cell") on the CIFAR-10 dataset and then this cell was later applied to the ImageNet dataset by iteratively stacking copies of this cells, each with their own set of hyperparameters resulting in a novel architecture (Fig. 1).

4.1.2 ResNET 50

The ResNet-50 [24] model is composed by 5 so called "stages" each composed by a convolution and an Identity block. Each convolution block and each identity block have 3 convolution layers which results in over 23 million trainable parameters. ResNET is theoretically important because it introduced two major breakthroughs in computer vision:
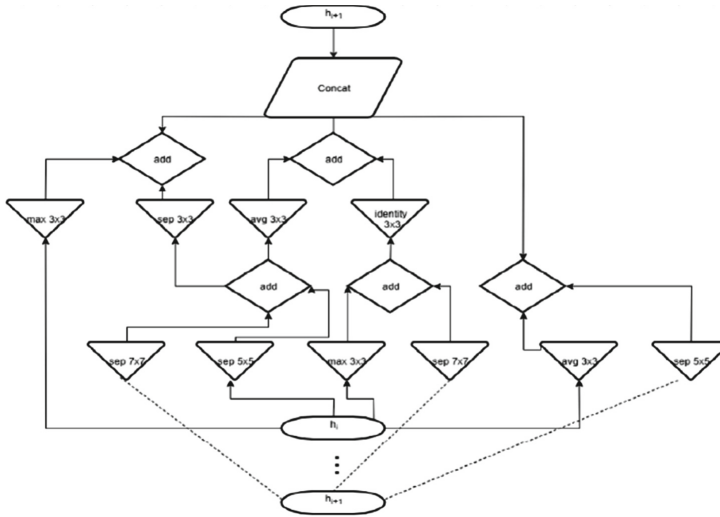
Fig. 1. NASNet large architecture

1. The mitigation of the gradient vanishing problem by allowing this alternate shortcut path for reinjecting information to the flow
2. The possibility to learn the identity function of the previous output, by ensuring that the later layers will perform at least as good as the previous (Fig. 2).

Fig. 2. ResNET-50

4.1.3 Inception V3

Inception-v3 [25] is the third release of a convolutional neural network architecture developed at Google which derived from the Inception family. This architecture makes several improvements including using Label Smoothing, factorized convolutions, batch normalization and auxiliary classifier which is used to propagate label information lower down the network (Fig. 3 and 4).
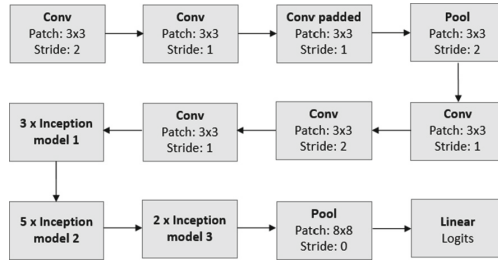
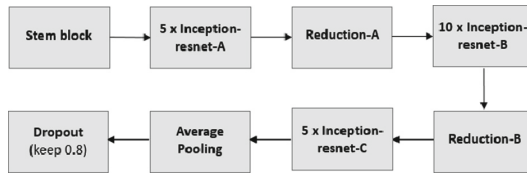**Fig. 3.** Diagram representation of inception V3 architecture

Fig. 4. Diagram representation of inception ResNet V2 architecture

4.1.4   Inception-ResNet-v2

Inception-ResNet-v2 [26] often called Inception V4 is a convolutional neural architecture that is built, as the name suggests, by fusing two major architecture families: Inception family e.g. Inception V3 and ResNet family by incorporating residual connections. This is at the moment one of the state of the art architecture used in image recognition tasks.

4.2   Bi-directional LSTM RNN for Online Recognition

For online recognition using recurrent neural networks, a novel Bi-Directional LSTM recurrent neural network is developed with the aim of performing online handwriting recognition. This online recognition is based solely on time series of x,y coordinates, no other information are provided. Thus, as in deep learning fashion, this architecture will automatically exploit long and short-term coherence and patterns with the aim of recognizing neurodegenerative diseases from just raw coordinates. Differently from Long-Short Term Memory RNN (briefly LSTM), bidirectional LSTM run the inputs in two ways: both from past to future and backward. This process preserves information from the future and from the past by combining the two hidden states (one for forward and one for backward) in order to preserve information from past and future. Authors in [27] have used bidirectional LSTM for modelling online handwriting recognition. The architecture developed also contains an Attention Mechanism layer. [28] The attention mechanism was invented for Natural Language Processing tasks where the encoder-decoder recurrent neural network architecture was used to learn to encode input sequences into a fixed-length internal representation, and second set of LSTMs read the internal representation and decode it into an output sequence. To overcome the problem that all input sequences are forced to be encoded into an internal vector of fixed length, a selective attention mechanism was developed with the aim to select these inputs and

relate them with respect to the output sequence. [28] This attention mechanism searches for a set of positions in the input where the most relevant information is concentrated. It does so by encoding the input vector into a sequence of vectors and then it adaptively chooses a subset of vectors while producing the output. [28] The intuition here is that attention mechanism would capture very long-term relations among coordinates in such a way to increase correlations among handwriting patterns of people affected by some neurodegenerative disease versus the normative sample. The architecture developed is depicted in Fig. 5 and was trained in an end-to-end fashion. It is composed by a bidirectional LSTM layer with 32 neurons followed by a dense layer with 32 neurons and ReLU activation function, this followed by an Attention layer with 32 neurons. At the end there is a dense layer with softmax activation function that carries out the classification.
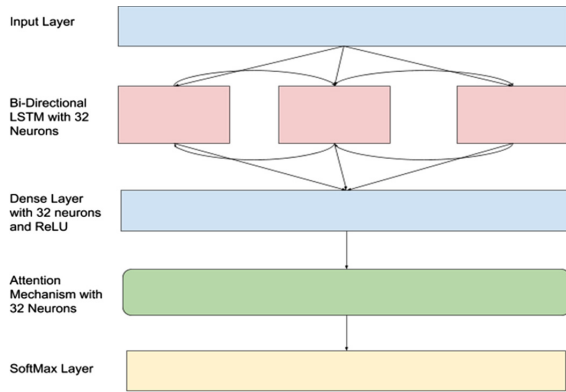

Fig. 5. Bidirectional LSTM with attention mechanism for online handwriting recognitionDiagram representation of Inception ResNet V2 architecture

5 Dataset Description and Results

5.1 Dataset Description

Raw data were collected by measuring x and y coordinates of the pen position and their timestamps. The pen inclination (tilt-x and tilt-y) and pressure of the pen's tip on the surface were also registered. Another important collected parameter was the "button status", i.e. a binary variable which gives 0 for pen-up state (in-air movement) and 1 for pen-down state (on-surface movement). A matrix X = (x, y, p, t, tilt_x, tilt_y, b) where each column is a vector of length N, where N is the number of sampled points, thus can describe the whole execution process of a single task. All the tasks are listed in Table 3.

The check copying task consists of asking the user to copy a check as shown in Fig. 6.

Another task is based on asking the user to find and mark a subset of predefined numbers inside matrices, as shown in Fig. 7.

The trail test consists of completing a succession of letters or numbers inside circles by linking them with other ones generating a path of variable complexity. The example in Fig. 8 is a clear example of a user affected by a neurodegenerative disease.

Table 3. Taks used

Task name	Task description
Chk	Check copying task
M1	Matrix 1
M2	Matrix 2
M3	Matrix 3
Tmt1	Trail 1 of connecting path
Tmt2	Trail 2 (difficult) of connecting path
Tmtt1	Trail test 1 of connecting path
Tmtt2	Trail test 2 (difficult) of connecting path

Fig. 6. Check copying task performed by a patient with some form of dementia
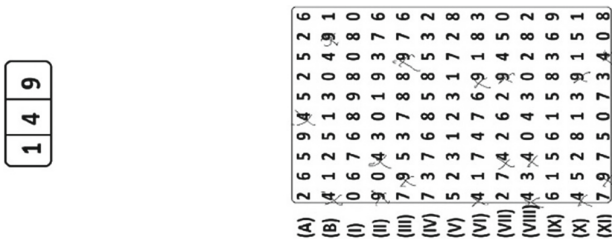

Fig. 7. M3 Matrix task

The user subset is composed by 42 subjects: 21 among them are affected by a neurodegenerative disease at different levels of severity which will be qualified as "mild", "assessed", "severe", "very severe". The other 21 are healthy control subjects. The dataset size is in line with sizes of other datasets mentioned in state of art review.

At this stage of the study, age and sex are not taken into account in the analysis. A deeper analysis won't be able to leave these parameters out of consideration.
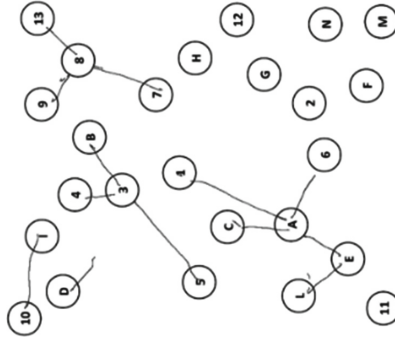
Fig. 8. Trail test number 2 mixing letters and numbers

5.2 Results

Table 4 shows the results. The accuracy is expressed as F1 score.

Table 4. Results of various techniques with respect to various tasks

Task name	NASNet large accuracy	ResNet 50 accuracy	Inception V3 accuracy	Inception ResNet v2 accuracy	Shallow learning	Bi directional LSTM w/attention
Check copying	0.72	0.72	0.80	0.78	**0.853**	0.72
M1	0.68	0.64	**0.76**	0.74	0.677	0.66
M2	0.65	0.58	0.65	0.69	**0.702**	0.69
M3	0.67	**0.75**	0.69	0.58	0.704	0.66
Tmt1	0.76	0.63	0.89	0.74	**0.799**	0.78
Tmt2	0.59	0.59	0.68	**0.74**	0.726	0.71
Tmtt1	0.84	0.68	0.64	0.74	0.774	**0.86**
Tmtt2	0.58	0.71	0.74	0.63	**0.858**	0.70
All tasks	0.66	0.70	0.67	0.72	**0.923**	0.74

6   Results discussion

In Table 4, different CNN ("NASNET LARGE", "RESNET 50", "INCEPTION V3", "Inception Resnet V2") and RNN architectures ("Bidirectional LSTM with Attention") were tested in order to understand their performances on detecting the presence (or absence) of a neurodegenerative disease by analyzing a series of previously described tasks (CHK, M1, M2, M3, TMT1, TMT2, TMTT1, TMTT2). Moreover, further analysis was performed by running the various techniques on a dataset obtained by merging data of all the tasks. In the following analysis positive class will be represented by 1

(those affected by neurodegenerative disease) and negative class by 0 (those without neurodegenerative disease). The most promising results were obtained using predefined features and doing the analysis based on the shallow learning approach, i.e. performing feature engineering by carefully selecting features from a set of physical parameters followed by automatic feature selection to decrease dimensionality: performances were characterized by a relatively small variance between different tasks, which suggests low dependency of accuracy from the specific task dataset. The second best outcome was from "Bidirectional LSTM" with attention, this deep recurrent neural network architecture achieved the lowest variability among accuracies at different tasks. Moreover, this network was capable of successfully exploiting neurodegenerative diseases patterns capable of binary discern healthy from un-healthy subjects based solely on the raw time series of x,y coordinates. All other deep learning neural networks trained on offline (static) images show significant variations with their accuracies from a task to another. This results in high variability of accuracies between tasks and thus a decrease of confidence. This analysis suggests that online handwriting outperforms the offline one both with a preliminary features selection or letting the algorithm to learn the most efficient patterns from raw data.

7   Conclusions

In this work, classic features have been employed for healthy/unhealthy binary classification of subjects included in the new dataset. The main goal of this work is to provide a benchmark of accuracy of different techniques available for neurodegenerative disease detection. Indeed, the analysis was performed on a specific subset of variables acquired during the handwriting tasks performance, specifically x, y coordinates and azimuth. The shallow learning approach, with a feature preselection, outperformed all the others architectures showing small variance of accuracies between different tasks. Similar results were obtained using "Bidirectional LSTM" with attention, while other deep learning algorithms were affected by higher variability in accuracy depending on the specific task analyzed. These results suggest that online handwriting is a better approach compared to the offline one, either with features preselection and with the algorithm learning itself from raw data. This last point opens new frontiers in automatic learning specific neurodegenerative disease patterns from timeseries of raw x,y coordinates. The next evolution of this work will be to perform not only binary prediction of healthy/unhealthy subjects but also to evaluate the severity level of diseases. In this regard, as the dataset is provided with multiple sessions of acquisitions for the same patients, it will also be analyzed the inferability of increments or decrements of disease severity with time, with respect to the adoption of medical treatments.

References

1. Impedovo, D., Pirlo, G.: Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. IEEE Rev. Biomed. Eng. **12**, 209–220 (2019)
2. De Stefano, C., Fontanella, F., Impedovo, D., Pirlo, G., di Freca, A.S.: Handwriting analysis to support neurodegenerative diseases diagnosis: a review. Pattern Recogn. Lett. **121**, 37–45 (2018)
3. Rosenblum, S., Samuel, M., Zlotnik, S., Erikh, I., Schlesinger, I.: Handwriting as an objective tool for Parkinson's disease diagnosis. J. Neurol. **260**(9), 2357–2361 (2013)
4. Astrom, F., Koker, R.: A parallel neural network approach to prediction of Parkinson's Disease. Expert Syst. Appl. **38**(10), 12470–12474 (2011)
5. O'Reilly, C., Plamondon, R.: Development of a sigma–lognormal representation for on-line signatures. Pattern Recogn. **42**(12), 3324–3337 (2009)
6. Pereira, C.R., et al.: A step towards the automated diagnosis of Parkinson's disease: analyzing handwriting movements. In: IEEE 28th International Symposium on Computer Based Medical Systems (CBMS), pp. 171–176 (2015)
7. Kahindo, C., El-Yacoubi, M.A., Garcia-Salicetti, S., Rigaud, A., Cristancho-Lacroix, V.: Characterizing early-stage alzheimer through spatiotemporal dynamics of handwriting. IEEE Signal Process. Lett. **25**(8), 1136–1140 (2018)
8. Caligiuri, M.P., Teulings, H.L., Filoteo, J.V., Song, D., Lohr, J.B.: Quantitative measurement of handwriting in the assessment of drug-induced Parkinsonism. Hum. Mov. Sci. **25**(4), 510–522 (2006)
9. Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., Faun-dez-Zanuy, M.: Decision support framework for Parkinson's disease based on novel handwriting markers. IEEE Trans. Neural Syst. Rehabil. Eng. **23**(3), 508–516 (2015)
10. Ponsen, M.M., Daffertshofer, A., Wolters, E.C., Beek, P.J., Berendse, H.W.: Impairment of complex upper limb motor function in de novo Parkinson's disease. Parkinsonism Related Disord. **14**(3), 199–204 (2008)
11. Smits, E.J., et al.: Standardized handwriting to assess bradykinesia, micrographia and tremor in Parkinson's disease. PLOS One **9**(5), e97614 (2014)
12. Broderick, M.P., Van Gemmert, A.W., Shill, H.A.: Hypometria and bradykinesia during drawing movements in individuals with Parkinson disease. Exp. Brain Res. **197**(3), 223–233 (2009)
13. Kotsavasiloglou, C., Kostikis, N., Hristu-Varsakelis, D., Arnaoutoglou, M.: Machine learning-based classification of simple drawing movements in Parkinson's disease. Biomed. Signal Process. Control **31**, 174–180 (2017)
14. Li, G., et al.: Temperature based restricted Boltzmann Machines. Sci. Rep. **6**, Article no. 19133 (2016)
15. Impedovo, D.: Velocity-based signal features for the assessment of Parkinsonian handwriting. IEEE Signal Process. Lett. **26**(4), 632–636 (2019)
16. Rao, K.R., Yip, P.: Discrete Cosine Transform: Algorithms, Advantages. Applications. Academic press, New York (2014)
17. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
18. Baraniuk, R.G.: Compressive sensing [lecture notes]. IEEE Signal Process. Mag. **24**, 118–121 (2007)
19. Reitan, R.M.: Validity of the Trail Making Test as an indicator of organic brain damage. Perceptual Motor Skills **8**(3), 271–276 (1958)
20. Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C.: A survey on deep transfer learning. In: International Conference on Artificial Neural Networks, pp. 270–279. Springer, Cham (2018)

21. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE, June 2009

22. Chollet, F.: Keras. Keras documentation: Keras Applications. Keras.io (2020). https://keras.io/api/applications/. Accessed 30 Oct 2020

23. Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V.: Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8697–8710 (2018)

24. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778 (2016)

25. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016)

26. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv preprint arXiv:1602.07261 (2016)

27. Liwicki, M., Graves, A., Fernàndez, S., Bunke, H., Schmidhuber, J.: In Proceedings of the 9th International Conference on Document Analysis and Recognition, ICDAR 2007 (2007)

28. Bahdanau, D., Cho, K., Bengio, Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473 (2014)

29. Pereira, C.R., Weber, S.A., Hook, C., Rosa, G.H., Papa, J.P.: Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics. In: 2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 340–346. IEEE, October 2016

30. Isenkul, M., Sakar, B., Kursun, O.: Improved spiral test using digitized graphics tablet for monitoring Parkinson's disease. In: Proceedings of the International Conference on e-Health and Telemedicine, pp. 171–175, May 2014

31. Impedovo, D., et al.: Writing generation model for health care neuromuscular system investigation. In: International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics, pp. 137–148. Springer, Cham, June 2013

32. Likforman-Sulem, L., Esposito, A., Faundez-Zanuy, M., Clémençon, S., Cordasco, G.: EMOTHAW: a novel database for emotional state recognition from handwriting and drawing. IEEE Trans. Hum. Mach. Syst. 47(2), 273–284 (2017)

33. Diaz, M., Ferrer, M.A., Impedovo, D., Pirlo, G., Vessio, G.: Dynamically enhanced static handwriting representation for Parkinson's disease detection. Pattern Recogn. Lett. 128, 204–210 (2019)

34. Impedovo, D., Pirlo, G., Vessio, G., Angelillo, M.T.: A handwriting-based protocol for assessing neurodegenerative dementia. Cognit. Comput. 11(4), 576–586 (2019)

35. Puigcerver, J.: Are multidimensional recurrent layers really necessary for handwritten text recognition? In: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), vol. 1, pp. 67–72. IEEE, November 2017

36. Doetsch, P., Zeyer, A., Ney, H.: Bidirectional decoder networks for attention-based end-to-end offline handwriting recognition. In: 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 361–366. IEEE, October 2016

37. Dentamaro, V., Impedovo, D., Pirlo, G.: Gait analysis for early neurodegenerative diseases classification through the Kinematic Theory of Rapid Human Movements. IEEE Access (2020)

38. Dentamaro, V., Impedovo, D., Pirlo, G.: Sit-to-stand test for neurodegenerative diseases video classification. In: International Conference on Pattern Recognition and Artificial Intelligence, pp. 596–609. Springer, Cham, October 2020