Handwriting analysis to support Alzheimer Disease diagnosis: a preliminary study *

Nicole Dalia Cilia, Claudio De Stefano, Francesco Fontanella, Mario Molinara, and Alessandra Scotto Di Freca

Department of Electrical and Information Engineering (DIEI)
University of Cassino and Southern Lazio
Via G. Di Biasio, 43 03043
Cassino (FR), Italy
(nicoledalia.cilia,destefano,fontanella,m.molinara,a.scotto)@unicas.it

Abstract. Alzheimers disease (AD) is the most common neurodegenerative dementia of old age and the leading chronic disease contributor to disability and dependence among older people worldwide. Handwriting is among the motor activities compromised by AD, which is the result of a complex network of cognitive, kinaesthetic and perceptive-motor skills. Indeed, researchers have shown that the patients affected by these diseases exhibit alterations in the spatial organization and poor control of movement. In this paper, we present the preliminary results of a study in which an experimental protocol (including the copy of words, letters and sentence task) has been used to assess assess the kinematic properties of the movements involved in the handwriting. The obtained results are very encouraging and seem to confirm the hypothesis that machine learning-based analysis of handwriting can be profitably used to support AD diagnosis.

Keywords: Handwriting · Classification algorithm · Alzheimer Disease

1 Introduction

Alzheimers disease (AD) is the most prevalent brain neurodegenerative disorder progressing to severe cognitive impairment and loss of autonomy (i.e., dementia) in older people [16, 10, 9]. Criteria for clinical diagnosis of AD were proposed in 1984 [5] by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and by the Alzheimers Disease and Related Disorders Association (ADRDA). According to these criteria, the diagnosis of AD needs histopathologic confirmation (i.e., microscopic examination of brain tissue) in autopsy or biopsy [1, ?,?]. Nonetheless, an early diagnosis would greatly improve the effectiveness of available treatments, but it is still a challenging task.

To date, clinical diagnosis of such diseases are performed by physicians and may be supported by tools such as imaging (e.g. magnetic resonance imaging),

* This work is supported by the Italian Ministry of Education, University and Research (MIUR) within the PRIN2015-HAND project.

blood tests and lumbar puncture (spinal tap). Recently, researchers have shown that the patients affected by these diseases exhibit alterations in the spatial organization and poor control of movements. It follows that, in principle, some diagnostic signs of AD should be detectable by motor tasks. Among them, Handwriting (HW), which is the result of a complex network of cognitive, kinaesthetic and perceptive-motor skills [15] may be significantly compromised. For example, in the clinical course of AD, dysgraphia occurs both during the initial phase, and in the progression of the disorder [6]. However, in this field, many published studies have been conducted in the areas of medicine and psychology, where typically standard statistical tools (with ANOVA and MANOVA analysis) are used to investigate the relationship between the disease and the variables taken into account to describe a patients handwriting. Conversely, as we have shown in [2], very few studies have been published, which use classification algorithms to detect people affected by AD from their handwriting.

From the literature reviewed, we found several open issues that need to be addressed: (i) the definition of an experimental protocol, composed of an ensemble of handwriting and drawing tasks the subjects should perform, to investigate whether there are specific features allowing us to detect the early signs of AD; (ii) a well-designed data set, large enough to allow an effective training of classification algorithms; (iii) the tools to be used for the automatic analysis of the handwriting tasks. To solve the first issue we proposed [**?**] a protocol consisting in 25 tasks (copy, reverse copy, free writing, drawing etc.) to analyse the impact of different tasks and different motor skills on AD patient performance.

In this paper, we present the results of a preliminary study in which we have considered only a subset of the tasks included in the above protocol, in order to assess the kinematic properties of the movements involved in the handwriting. We collected the data produced by 130 subjects using a graphic tablet. From these data, we have extracted the most common features in literature [2]. Finally, for the classification phase, we considered two effective and widely used classification methods, namely Random Forest and Decision Trees. The obtained results are very encouraging and seem to confirm the hypothesis that machine learning-based analysis of handwriting can be profitably used to support the diagnosis of AD.

2 Materials and Methods

In the following subsections, the dataset collection procedure and the protocol designed for collecting handwriting samples, are detailed.

2.1 Dataset collection

The 130 subjects who participated to the experiments, namely 66 AD patients and 64 healthy controls, were recruited with the support of the geriatric ward, Alzheimer unit, of the "Federico II" hospital in Naples. As concerns the recruiting criteria, we took into account clinical tests (such as PET, TAC and enzymatic analyses) and standard cognitive tests (such as MMSE). In these tests, the cognitive abilities of the examined subject were assessed by using questionnaires including questions and problems in many areas, which range from orientation to time and place, to registration recall. As for the healthy controls, in order to have a fair comparison, demographic as well as educational characteristics were considered and matched with the patient group. Finally, for both patients and controls, it was necessary to check whether they were on therapy or not, excluding those who used psychotropic drugs or any other drug that could influence their cognitive abilities.

The data were collected by using a graphic tablet, which allows the recording of pen movements during the handwriting process. During the trial, images and sound stimuli are also provided to the subject to guide the execution of the tasks. Finally, the subjects were also asked to follow the indications provided by the experimenter.

2.2 Protocol

The aim of the protocol is to record the dynamics of the handwriting, in order to investigate whether there are specific features that allow us to distinguish subjects affected by the above-mentioned diseases from healthy ones. The nine tasks considered for this study are selected from a larger experimental protocol presented in [**?**], and they are arranged in increasing order of difficulty, in terms of the cognitive functions required. The goal of these tasks is to test the patients' abilities in repeating complex graphic gestures, which have a semantic meaning, such as letters and words of different lengths and with different spatial organizations. The tasks have been selected according to the literature, which suggests that:

(i) graphical tasks and free spaces allow the assessment of the spatial organization skills of the patient;
(ii) the copy and dictation tasks allow to compare the variations of the writing respect to different stimuli (visual or sound);
(iii) tasks involving different pen-ups allow the analysis of air movements, which it is known to be altered in the AD patients;
(iv) tasks involving different graphic arrangements, e.g. words with ascenders and/or descendants, or complex graphic shapes, allow testing fine motor control capabilities.

Furthermore, in order to evaluate patient responses under different fatigue conditions, these tasks should be provided by varying their intensity and duration.

(1) As in [17] or in [11], in the first task the subjects must copy three letters that have different graphic composition.
(2) The second task consists in copying four letters on adjacent rows. The aim of the cues is to test the spatial organization abilities of the subject [13].

(3-4) The task 3 and 4 require participants to write continuously for four times, in cursive, a single letter and a bigram, respectively [14, 8]. These tasks allow the testing of the motion control alternation.
(5-8) The tasks 5, 6, 7 and 8 implies word copying, which is the most explored activity in the analysis of handwriting for individuals with cognitive impairment [12, 17, 8]. Moreover, to observe the variation of the spatial organization, we have introduced the copy of the same word without or with a cue.
(9) In the ninth task, subjects are asked to write, above a line (the cue), a simple phrase, dictated them by the experimenter. As in [7], the hypothesis is that the movements can be modified because of the lack of visualization of the stimulus.

2.3 Segmentation and feature extraction

The features extracted during the handwriting process have been exploited to investigate the presence of neurodegenerative diseases in the examined subjects. We used the MovAlyzer tool to process the handwritten trace, considering both on-paper and on-air traits. Their segmentation in elementary strokes was obtained by assuming as segmentation points both pen up and pen down, as well as the zero crossings of the vertical velocity profile. The feature values were computed for each stroke and averaged over all the strokes relative to a single task: we considered for each feature both the mean value and the maximum value for that task. Note that, as suggested in [17], we have separately computed the features over on-paper and on-air traits, because the literature shows significant differences in motor performance in these two conditions.

3 Experiments and results

The software used for the classification was Weka. We decided to use three data groups: the data related to the on-air features; the data concerning the on paper features; and the overall data (on paper and on air). Each feature is present in the data set for the 9 tasks performed by the 130 subjects. For the experiments, we used two different classifiers: The Random Forest and the Decision Tree, namely J48. For both of them, 500 iterations were performed and a 5 fold validation strategy was considered. It is noteworthy that we have chosen the Random Forest classifier because it is widely recognized as a top of performing classifier. However, being an ensemble of classifiers, it does not provide easily interpretable models. We also have chosen Decision Tree because through the generated tree it is possible to identify the specific features used during the classification process.

The tables below summarize the values of Accuracy and False Negative Rate (FNR) for each task. The first column provides the number of tasks, the second the features used, the third the classifier employed, while in the fourth we report

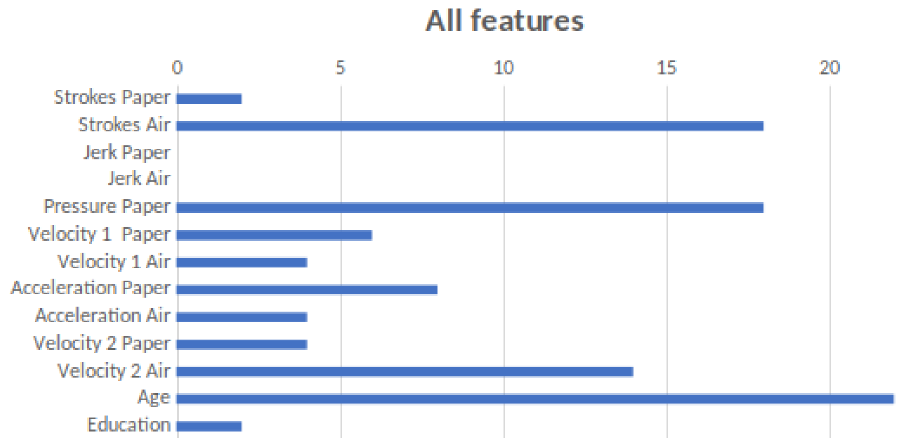


Fig. 1. Occurrences of all features.

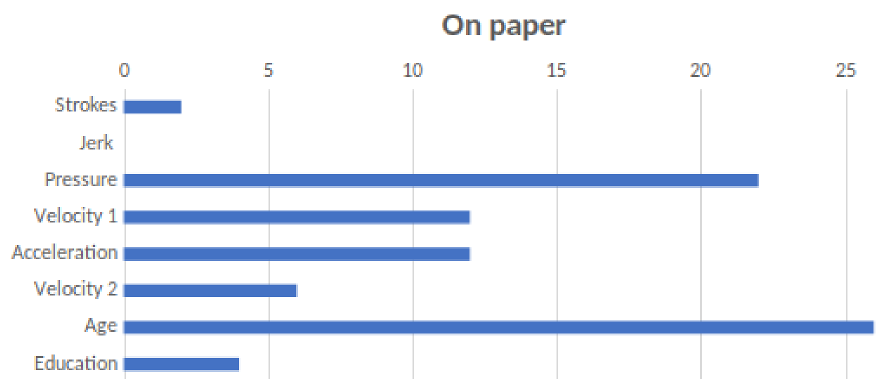
the value of accuracy and in the final column, we show the value of False Negative Rate. The False Negative Rate is very relevant in medical diagnosis applications since it indicates the ability of correctly identifying the patients, thus allowing their inclusion in the appropriate therapeutic pathway.

The preliminary results seem to encourage the use of classification algorithms as tools to support the diagnose of AD. From the tables shown below (Tables 1-4) we can point out that: firstly, for each task the max value (in bold) of accuracy is over 70 %, reaching peak in some tasks, such as the fifth task with value of 76 %. Secondly we can claim that, on average, emerges a better classification using the Random Forest classifier compared to J48. This is easily justifiable considering that Random Forest, unlike J48, is an ensemble of classifiers. However, as reported in the last column, FNR is lower using J48 classifier. In particular, the lower value of FNR occurs in the on-paper traits of the second task, with a value of 8.82 %. Moreover, as shown in the histograms (Fig. 2 and 3), from a decision tree it is possible to identify the occurrences of the features for the classification. It is noteworthy that, using all features, the system mainly uses the strokes, the pressure and the subject age that is the most used one. Moreover, jerk, both on air and on paper, is never used. Instead, if we consider the two groups of features separately, subject age still remains the most used one for both groups, followed by the velocity (on air) and pressure (on paper).

4 Conclusion and open issues

In this paper, we presented a novel solution for the early diagnosis of Alzheimer's disease by analyzing features extracted from handwriting. The preliminary results obtained are encouraging and the work is in progress to increase general performance. In particular, from the results obtained we can claim to have iden-

(a)



(b)

Fig. 2. Occurrence of features. On air (a) and on paper (b).

Table 1. Classification results of task 1 and 2.

Task	Features	Classifier	Accuracy	FNR
1	All	RF	71.96	28.79
	All	J48	66.66	19.70
	On paper	RF	**72.72**	28.79
	On paper	J48	65.90	27.27
	On air	RF	71.21	22.73
	On air	J48	68.18	18.18
2	All	RF	66.41	33.82
	All	J48	66.41	36.76
	On paper	RF	**72.38**	23.53
	On paper	J48	67.16	36.76
	On air	RF	60.44	33.82
	On air	J48	70.89	**8.82**

tified some simple tasks, features and classifiers, which support the diagnosis of Alzheimer's. In other words, this system could represent a low cost and non-

Table 2. Classification results of task 3 and 4.

Task	Features	Classifier	Accuracy	FNR
3	All	RF	**69.09**	44.90
	All	J48	61.81	40.82
	On paper	RF	68.18	44.90
	On paper	J48	63.63	30.61
	On air	RF	67.27	44.90
	On air	J48	66.36	22.45
4	All	RF	**71.42**	36.00
	All	J48	63.81	30.00
	On paper	RF	66.66	38.00
	On paper	J48	67.62	34.00
	On air	RF	61.90	42.00
	On air	J48	63.80	18.00

invasive tool to support the diagnosis of Alzheimer's disease, in juxtaposition with actual diagnosis systems.

Starting from these results, the next steps to be taken could include the combination of all tasks taken into account in a suitable way (the task itself can be used as a new feature) [4, 3]; introduction of a reject option in order to reduce false negative rate; introduction of new features evaluated on slant, loop surface, horizontal size and vertical size, etc; reduction of the unbalancing dataset caused by difficulties in recruiting young patients and old people without any cognitive disease.

References

1. C, B., A, T., A, L., S, C., G, T., V, B., A, S., R, F., F, N., et al.: Classification of single normal and alzheimer's disease individuals from cortical sources of resting state eeg rhythms. Frontiers in Neuroscience 10(61)(61), 1–18 (2016)
2. C, D.S., F, F., D, I., G, P., di Freca A, S.: Handwriting analysis to support neurodegenerative diseases diagnosis: a review. Pattern Recognition Letters 121, 37–45 (2018)
3. Cordella, L., De Stefano, C., Fontanella, F., Scotto Di Freca, A.: A weighted majority vote strategy using bayesian networks. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8157 LNCS(PART 2), 219–228 (2013)
4. De Stefano, C., Fontanella, F., Folino, G., Scotto Di Freca, A.: A bayesian approach for combining ensembles of gp classifiers. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 6713 LNCS, 26–35 (2011)

Table 3. Classification results of task 5, 6, 7, 8.

Task	Features	Classifier	Accuracy	FNR
5	All	RF	75.67	25.45
	All	J48	66.66	38.18
	On paper	RF	**76.57**	23.64
	On paper	J48	67.56	41.82
	On air	RF	64.68	34.55
	On air	J48	61.26	38.18
6	All	RF	69.29	39.22
	All	J48	64.91	41.18
	On paper	RF	68.41	43.14
	On paper	J48	**71.93**	21.57
	On air	RF	64.03	45.10
	On air	J48	63.15	39.22
7	All	RF	63.63	38.46
	All	J48	53.63	59.62
	On paper	RF	**64.54**	38.46
	On paper	J48	62.72	40.38
	On air	RF	63.63	38.46
	On air	J48	62.72	21.15
8	All	RF	70.43	36.54
	All	J48	69.56	38.46
	On paper	RF	68.69	38.46
	On paper	J48	66.95	38.46
	On air	RF	**72.17**	30.77
	On air	J48	67.82	34.62

Table 4. Classification results of task 9.

Task	Features	Classifier	Accuracy	FNR
9	All	RF	**72.30**	34.85
	All	J48	66.92	31.75
	On paper	RF	70.00	26.98
	On paper	J48	67.69	12.70
	On air	RF	70.00	34.92
	On air	J48	69.23	33.33

5. G, M., D, D., M, F., R, K., D, P., M, S.E.: Clinical diagnosis of alzheimers disease: report of the nincdsadrda work group under the auspices of department of health

and human services task force on alzheimers disease. Neurology 34, 939–944 (1984)
6. H, Y.J., S, R., P, M., RS, D., H, L.: Alzheimers disease and mild cognitive impairment deteriorate fine movement control. Journal of Psychiatric Research 42, 1203–1212 (2008)
7. Hayashi, A., Nomura, H., Mochizuki, R., Ohnuma, A., Kimpara, T., Ootomo, K., Hosokai, Y., Ishioka, T., Suzuki, K., Morio, E.: Neural substrates for writing impairments in japanese patients with mild alzheimer's disease: A spect study. Neuropsychologia 49(7), 1962–1968 (2011)
8. Impedovo, D., Pirlo, G.: Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. IEEE Reviews in Biomedical Engineering pp. 1–13 (2018)
9. J, K., G, L.H., A, U., M, S.J., L, M.C., H, G.K., et al.: The precursor of alzheimers disease amyloid a4 protein resembles a cell-surface receptor. Nature 325, 733–736 (2015)
10. L, P.D.: Aging of the brain and dementia of the alzheimer type. Princ. neural Sci. p. 11491168 (2000)
11. Lambert, J., Giffard, B., Nore, F., de la Sayette, V., Pasquier, F., Eustache, F.: Central and peripheral agraphia in alzheimer's disease: From the case of auguste d. to a cognitive neuropsychology approach. Cortex 43(7), 935–951 (2007)
12. Onofri, E., Mercuri, M., Archer, T., Ricciardi, M.R., F.Massoni, Ricci, S.: Effect of cognitive fluctuation on handwriting in alzheimer's patient: A case study. Acta Medica Mediterranea 3, 751 (2015)
13. Onofri, E., Mercuri, M., Salesi, M., Ricciardi, M., Archer, T.: Dysgraphia in relation to cognitive performance in patients with Alzheimer's disease. Journal of Intellectual Disability-Diagnosis and Treatment 1, 113–124 (2013)
14. Slavin, M.J., Phillips, J.G., Bradshaw, J.L., Hall, K.A., Presnell, I.: Consistency of handwriting movements in dementia of the alzheimer's type: a comparison with huntington's and parkinson's diseases. J Int Neuropsychol Soc. 5(1), 20–25 (1999)
15. Tseng, M.H., Cermak, S.A.: The influence of ergonomic factors and perceptual–motor abilities on handwriting performance. American Journal of Occupational Therapy 47(10), 919–926 (1993)
16. V, J., T, D., K, A., O, A., B, W., A, N., N, T.: Longitudinal changes in quantitative eeg during long-term tacrine treatment of patients with alzheimers disease. Neuroscience Letters 254(4), 8588 (1988)
17. Werner, P., Rosenblum, S., Bar-On, G., Heinik, J., Korczyn, A.: Handwriting process variables discriminating mild alzheimer's disease and mild cognitive impairment. Journal of Gerontology: PSYCHOLOGICAL SCIENCES 61(4), 228–36 (2006)