Lognormal Features for Early Diagnosis of Alzheimer's Disease Through Handwriting Analysis

Nicole Dalia Cilia[2,4], Tiziana D'Alessandro[3], Cristina Carmona-Duarte[1],
Claudio De Stefano[3], Moises Diaz[1], Miguel Ferrer[1],
and Francesco Fontanella[3(✉)]

[1] IDeTIC: Instituto para el Desarrollo Tecnológico y la Innovación en
Comunicaciones, Universidad de Las Palmas de Gran Canaria, Las Palmas, Spain
[2] Department of Computer Engineering, University of Enna "Kore", Cittadella
Universitaria, Enna, Italy
[3] Department of Electrical and Information Engineering (DIEI), University of
Cassino and Southern Lazio, Via G. Di Biasio 43, 03043 Cassino, FR, Italy
fontanella@unicas.it
[4] Institute for Computing and Information Sciences, Radboud University,
Toernooiveld 212, 6525 EC Nijmegen, The Netherlands

Abstract. Alzheimer's disease causes most of dementia cases. Although
currently there is no cure for this disease, predicting the cognitive decline
of people at the first stage of the disease allows clinicians to alleviate
its burden. Clinicians evaluate individuals' cognitive decline by using
neuropsychological tests consisting of different sections, each devoted
to test a specific set of cognitive skills. The sigma-lognormal model
allows complex movements to be represented as a summation of sim-
ple time-overlapped movements, and has been used in several fields to
model numerous human movements such as, for example, handwriting
and speech. Recently, this theory has been also used for detecting and
monitoring neurodegenerative disorders. In this paper, we present the
results of a preliminary study aimed at exploring the use of lognormal
features to classify patients affected by Alzheimer's disease. The promis-
ing results achieved confirms that lognormal features can be used to
support Alzheimer's diagnosis.

1 Introduction

Neurodegenerative diseases (NDs in the following) are incurable and debilitating,
caused by progressive degeneration of nerve cells, affecting movements and/or
mental skills. Alzheimer's disease (AD) is the most common among them, and
because of worldwide lifespan lengthening, it is expected that its incidence will
dramatically increase in the coming decades.

AD produces a slow and progressive decline in mental functions such as
memory, thought, judgment, and learning abilities. The predominant symptom

in the early stages of AD is the episodic memory impairment, whereas later stages are characterized by progressive amnesia and deterioration in other cognitive domains.

Unfortunately, there is no cure for AD, but its symptoms can be managed during their progression. This creates a critical need for the improvement of the approaches currently used for diagnosing them as early as possible. As cognitive and motor functions are both involved in planning and execution of movements, and because handwriting requires a precise and properly coordinated control of the body [20], the analysis of handwriting dynamics might provide a cheap and non-invasive method for evaluating the disease progression [10]. Furthermore, it has been observed that the application of machine learning methods to motor function has shown promise in decreasing the time taken to perform clinical assessments [1,12]. To this aim, cheap and widely used graphic tablets can be used to administer handwriting tests, which include simple and easy-to-perform handwriting/drawing tasks [10], and to record kinematic and dynamic information of the performed movements. For this reason, researchers are showing an increasing interest in developing and using machine learning based methodologies to support both the diagnosis and the treatment of NDs, and several methods have been proposed for the diagnosis of both AD [24].

The Kinematic Theory of rapid movements, together with the use of the Sigma-Lognormal model, allows the decomposition of a complex movement into a vector summation of simple time-overlapped movements [15–17]. This theory has been applied in several fields to model numerous human movements such as, handwriting [14], speech [2], head and trunk movement [11], etc. However, it has been barely applied to the detection and monitoring of neuromuscular disorders [13,19]. Specifically, this model has been used to classify parkinsonian patients in this pair of papers [8,9]. The authors found competitive performance by combining this model with other velocity-based features like Maxwell-Boltzmann distribution, Fourier, and Cepstrum transforms.

In this paper, we present the results of a preliminary study aimed at exploring the use of lognormal features to classify patients affected by AD, on the basis of their ability to accomplish six handwriting tasks. Those tasks were introduced in [4], and are described later in this paper. We collected the data produced by 174 participants (89 AD patients and 85 healthy people). To the best of our knowledge, this is the largest dataset containing handwriting data related to AD. Starting from the lognormal parameters computed to represent the handwriting contained in this data, we have identified fourteen features that can be used to characterize the handwriting of people affected by AD. We assessed the effectiveness of the features extracted We used seven well-known and widely used classifiers to asses the effectiveness of the features proposed. The promising results achieved confirms that lognormal features can be used to support AD diagnosis.

The organization of this paper is as follows. Section describes the Sigma-lognormal model used for the representation of handwriting. In Sect. 3 we present the tasks used to collect handwriting data and the features extracted using the

Sigma-lognormal model. Section 4 details some experimental results. Concluding remarks and possible future investigations are outolined in Sect. 5.

2   The Sigma-Lognormal Model

Based on lognormal movement decomposition, there are several studies about the normative range of variations in the lognormal parameters, which give a notion of how ideal a movement could be [18]. To parametrizing the human movement velocity and trajectory by the Kinematic Theory, different algorithms have been developed, as Robust XZERO [5,14] and IDeLog [6]. In this work we based on the IDeLog algorithm [6].

Sigma-Lognormal model considers the resulting velocity of each simple fast movement primitive as a lognormal function (Λ), being each peak of velocity between two speed minima modeling by a lognormal. The lognormal parameters, t_{0_j}, μ_j and σ_j^2 are calculated finding the less minimimun error between the velocity profile and the obtained lognormal from successive interactions and the trajectory original profile and the reconstructed one. The lognornormal function that model each velocity peak or "simple movement" or "stroke" can be defined as:

$$v_j(t; t_{0_j}, \mu_j, \sigma_j^2) = D_j \Lambda(t; t_{0_j}, \mu_j, \sigma_j^2) = \frac{D_j}{\sigma_j \sqrt{2\pi}(t - t_{0_j})} exp\{\frac{[-ln(t - t_{0_j}) - \mu_j]^2}{2\sigma_j^2}$$

(1)

where t is the time basis, D_j the amplitude , t_{o_j} the time of occurrence, μ_j the time delay and σ_j the response time, both on a logarithmic time scale.

In case a complex movement, a succession of simple movements or strokes as can be observed in Fig. 1, the velocity profile $v_n(t)$ is given by the time superposition of the M previous lognormals.

$$v_n(t) = \sum_{j=1}^{M} v_j(t; t_{0_j}, \mu_j, \sigma_j^2) = \sum_{j=1}^{M} D_j \begin{bmatrix} \cos(\Phi_j(t)) \\ \sin(\Phi_j(t)) \end{bmatrix} \Lambda(t; t_{0_j}, \mu_j, \sigma_j^2)$$
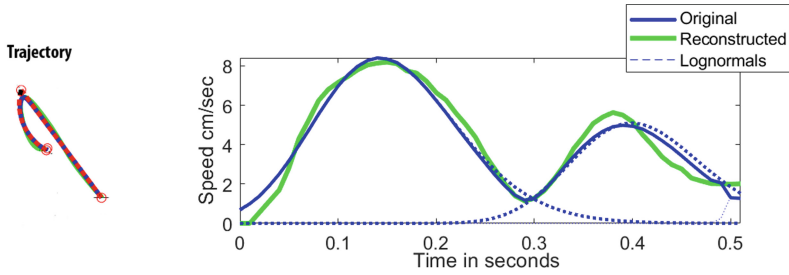
(2)

where $\Phi_j(t)$ is the angular position given by:

$$\Phi_j(t) = \Theta_{s_j} + \frac{(\Theta_{e_j} - \Theta_{s_j})}{2}[1 + erf(\frac{ln(t - t_{0_j}) - \mu_j}{\sigma_j \sqrt{2}})]$$

(3)

being Θ_{s_j} and Θ_{ej} are the starting and the end angular direction of the jth simple movement or stroke.

3   Tasks and Features

Following subsections detail the data collection procedure, the tasks used, and the features extracted.

Fig. 1. An example of lognormal.

3.1 Data Collection

Nowadays it is known that alteration in handwriting is one of the first signs of AD, that is why data acquisition step for this work focus on the recording and collection of handwriting samples. Those samples comes from the execution of a protocol [4] composed of different kinds of handwriting tasks. Every participant executed all the tasks with a special pen on A4 paper sheets fixed to a graphic tablet, that allow the recording of the handwriting in terms of x-y-z coordinates for each point, acquired at a constant sampling rate, equal to 200 Hz. The first two coordinates are spatial ones and represent the point position in the two-dimensional surface where the writing is produced, while the third one is a measure of the pressure exerted by the subject at that point. This last measure assumes positive values when the pen is on the sheet, while a null value when it is detached, up to a maximum distance of 3 cm from the sheet, beyond which the system is not able to receive information. The protocol was administered to a group of 174 participants: 89 patients at the first stages of AD and a control group of 85 people. Both the AD patients and the control group were recruited with the support of the geriatrics department, Alzheimer's unit, of the "Federico II" hospital in Naples. Both groups were selected according to a recruiting criteria based on standard clinical tests, such as the Mini-Mental State Examination (MMSE), the Frontal Assessment Battery (FAB) and the Montreal Cognitive Assessment (MoCA).

3.2 Tasks

In this study we considered only the handwriting samples relative to six tasks of the protocol:

1. Join two points with a vertical line continuously for four times. The up-down vertical movements require the finger joint and wrist movements. This task is useful to investigate elementary motor functions [27];
2. Trace a circle continuously for four times. The circle diameter has to be 6 cm. This task allows to test the automaticity of movements and the regularity and coordination of the sequence of movements [21];

3. Write continuously for four times, in cursive, the bigram 'le'. These letters allow to test the motion control alternation;
4. Copy in reverse order a simple italian word: "bottiglia" (bottle in English). This task has been inspired by the MMSE test, where one of the task requires people spelling a word backward;
5. In the fifth task a telephone number (10 digits) has to be written under dictation. The hypothesis underlying the introduction of this task is that motor planning in writing a telephone number is different from that for writing a word;
6. The sixth task is the Clock Drawing Test (CDT). In [26] the authors found that CDT shows a high sensitivity for mild AD.

The first two tasks belong to the category of graphic tasks, whose objective is to test the patient's ability in: (i) writing elementary traits; (ii) joining some points; (iii) drawing figures (simple or complex and scaled in various dimensions). The third and the fourth tasks are copy and reverse Copy tasks, whose objective is to test the person's abilities in repeating complex graphic gestures, which have a semantic meaning, such as letters, words and numbers (of different lengths and with different spatial organizations). The fifth is a dictation task, whose purpose is to investigate how the writing varies (with phrases or numbers) in which the use of the working memory is necessary throughout the execution. The sixth task is a graphic task whose purpose is not only to test the dynamic ability of a person, but also his cognitive skills, the spatial dysfunction and lack of attention. This test requires verbal understanding, memory and spatially coded knowledge in addition to constructive skills [25].

3.3   Lognormal Features

The feature engineering process allowed us to identify a set of features that according to our domain of knowledge were good candidates to discriminate the handwriting of people affected by AD from that of healthy people. The Sigma-Lognormal model, defined in Sect. 2 was applied to the data acquired as stated in Sect. 3.1. The result of this procedure was the decomposition of each task into a vector summation of simple time-overlapped movements, from which it was possible to extract a set of Sigma-Lognormal parameters $P_j = [D_j, t_{0_j}, \mu_j, \sigma_j, \Theta_{s_j}, \Theta_{e_j}]$. In particular, for every point (x, y) acquired during the execution of the tasks, one or more overlapping lognormals were found, so their parameters and the percentage of contribution were stored for every point. The term "First lognormal" is used to refer to the lognormal that most contributes for a certain point. Once the Sigma-Lognormal parameters were obtained for every task and every participant, it was possible to compute a set of fourteen features, described in Table 1.

Table 1. Summary of computed features

Name	Description
num_seg	Total number of segments generated by the execution of the task
avg_log	Average of the number of overlapped lognormals for every point
tot_log	Total number of lognormals extracted from the entire trace of the task
avg_D	Average of D parameter of the first lognormal for every point
D_max	Max of D parameter found among the first lognormals of all the points
P_first_log	Average of the percentage of contribution of the first lognormal for all the points
σ_stability	Variance of the sigma parameter of the first lognormal for all the points
diff_logs	Average of the differences between the percentage of contribution of the first and the second lognormal on all points
var_log	Variance of the percentage of contribution of the first lognormal on all points
avg_t_o	Average of the t_o parameter of the first lognormal on all points
avg_σ	Average of the σ parameter of the first lognormal on all points
avg_μ	Average of the μ parameter of the first lognormal on all points
avg_Θ_s	Average of the Θ_s parameter of the first lognormal on all points
avg_Θ_e	Average of the Θ_e parameter of the first lognormal on all points

The aim of this procedure is to use those computed features to distinguish between patients and healthy controls, the two groups of participants involved. From this section on, those features will be referred as "Lognormal Features".

4   Experimental Results

This section shows the results obtained by applying several classification approaches according to the input data. Specifically, lognormal features are classified through six well known ML algorithms, while RGB images are used to feed three different kinds of CNNs.

We used the lognormal features (see Sect. 3.3) with standard machine learning algorithms: k-Nearest Neighbors (K-NN), Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression (LR), Gradient Boosting (GB), XGboost (XGB). We used the scikit-Learn library. The settings of their hyperparameters were left at the default values provided by scikit-Learn. The only exceptions regard the SVM classifiers, for which we used a linear kernel, and the KNN classifier, for which a the number of neighbours was set to 3. In order to obtain statistically significant results, we performed 30 runs for each classifier. For each run, the dataset was randomly shuffled and a 5-fold classification strategy was adopted. In order to evaluate the performance of the mentioned models we considered the following metrics: accuracy (acc), Sensitivity (True Positive Rate, TPR), Specificity (True Negative Rate, TNR), Precision, False Negative Rate (FNR), and Area Under the Curve (AUC).

Since we performed 30 runs for each classifier, the above mentioned parameters were computed for each run and their average and standard deviation (in parentheses) are shown in the following tables. All the metrics are expressed in percentages, except for the AUC and bold values highlight the best performance achieved.

Looking at the accuracies in Table 2 it is worth noting that we achieved the best performance on task 3 with a value of accuracy equal to 74.66% (SVM), whereas the worst performance was obtained on the task 1 with an accuracy of 58.24% (DT). The best performing algorithm was SVM, except for tasks 2 and 4 where RF achieved higher values. The DT classifier, on the contrary, achieved the worst performances for almost every task.

Looking at the table, we can observe a general trend: the first two tasks have worse performances than the others. An explanation to this phenomenon can comes from the analysis of the considered tasks. As mentioned in Sect. 3 the first two tasks are graphic and test the dynamics of simple movements and the motor control of the person who executes them, without requiring an important cognitive attention. The other tasks are words, numbers and the clock drawing test and they indeed require cognitive attention, as some of them have semantic meanings, include descending and ascending traits, requiring greater coordination, control skills and the use of the working memory. These considerations suggest that the use of the lognormal features is more effective on tasks with a semantic meaning instead of graphic tasks and it better brings out the difference between patients and healthy controls.

Table 3 shows the sensitivity values obtained during the experimental process. The sensitivity is a very important metric to consider when facing problems in the medical field, as it gives information about the number of patients correctly recognized. The best sensitivity score is obtained by RF on the fourth task (77.47%), while the worst by DT on the first task (59.29%). According to this table, RF and LR classifiers achieved good sensitivity values, but this doesn't mean they are the best classifiers, because looking at the accuracy the best is SVM. Despite SVM is the best classifier, this table shows that other classifiers are better able to recognize patients correctly.

From Table 4 we can see the specificity values obtained. The best specificity measure is achieved by SVM on the third task (82.03%), while the worst by RF on the second task (54.06%). This measure is linked to the sensitivity as it gives information about the healthy control participants correctly classified. As we said SVM was the best classifier according to the accuracy, but didn't achieve the higher sensitivity values, the specificity table, as a consequence, show high values of this metric for the SVM classifier. It means SVM is the best classifier according the accuracy, but, taking into account the considerations on sensitivity and specificity, it seems to better recognize healthy controls instead of patients, among our participants.

Table 5 shows that the best precision value is achieved from SVM on the third task (80.36%), while the worst by DT on the first task (59.89%). Though according to the sensitivity SVM wasn't the better classifier in recognising patients, this table shows that it is the most precise.

The FNR values obtained during the experimental process are shown in Table 6, where we can see that the best value is obtained with RF on the fourth task (22.52%), while the worst by DT on the first task (40.71%). FNR is a metric linked to Sensitivity, in fact they are complementary measures. FNR represents the number of patients that are erroneously classified and of course it should be at the lowest possible value. This is a fundamental information in the medical field, because an error in classifying a patient is a more serious problem than an error on a healthy person.

Table 7 shows the AUC values. AUC measures the area under the ROC curve, which illustrates the diagnostic ability of a binary classifier as its discrimination threshold is varied with the higher value the better. From the table we can observe that LR on the third task achieved the best result (0.83), whereas the worst was obtained by DT on the first task (0.58).

Table 2. Average Accuracy achieved on 30 runs for every ML algorithm on lognormal features

T	Accuracy						
	KNN	RF	DT	SVM	LR	GB	XGB
1	64.3 (2.7)	63.8 (2.2)	58.2 (3.1)	**66.9** (2.2)	63.9 (1.5)	61.5 (2.3)	61.3 (2.9)
2	62.7 (2.4)	**63.9** (2.6)	60.8 (3,2)	59.4 (2.0)	61.3 (2.1)	63.3 (3.7)	62.0 (3.0)
3	62.7 (2.0)	72.9 (2.2)	67.0 (2.7)	**74.6** (1.5)	74.2 (1.5)	69.7 (2.7)	71.2 (2.4)
4	64.6 (2.6)	**72.0** (2.0)	62.1 (3.5)	68.7 (1.8)	70.4 (1.4)	69.9 (3.5)	70.5 (2.7)
5	66.8 (2.3)	71.8 (1.8)	63.5 (2.8)	**73.6** (1.9)	72.6 (1.9)	69.7 (3.2)	69.6 (2.9)
6	67.0 (2.8)	67.1 (3.0)	61.1 (3.9)	**72.6** (2.8)	71.5 (2.3)	68.4 (2.9)	70.1 (3.0)

Table 3. Average Sensitivity achieved on 30 runs for every ML algorithm on lognormal features

T	Sensitivity						
	KNN	RF	DT	SVM	LR	GB	XGB
1	67.2 (2.6)	64.3 (3.0)	59.2 (4.2)	64.5 (2.7)	**67.3** (2.1)	62.6 (2.5)	62.5 (3.1)
2	66.6 (3.5)	**72.0** (3.9)	63.5 (4.2)	62.5 (3.5)	64.5 (3.2)	70.5 (4.3)	67.9 (3.6)
3	63.5 (2.0)	68.0 (3.7)	66.3 (4.1)	67.3 (1.8)	**70.1** (2.0)	66.3 (4.2)	68.6 (3.5)
4	67.7 (3.0)	**77.4** (2.1)	66.1 (5.4)	68.8 (3.0)	73.4 (2.7)	75.5 (3.5)	75.9 (3.1)
5	62.4 (3.1)	**70.1** (3.0)	62.3 (4.1)	68.6 (2.8)	68.5 (2.1)	67.6 (4.2)	68.2 (4.3)
6	70.5 (3.5)	69.6 (4.6)	62.0 (6.2)	75.1 (3.4)	**75.2** (3.6)	71.1 (4.8)	72.5 (5.3)

4.1 Comparison Findings

To test the effectiveness of the lognormal features extracted, we compared the results shown in the above tables with those achieved by some deep neural networks (DL) trained on synthetic images generated from the raw data described

Table 4. Average Specificity achieved on 30 runs for every ML algorithm on lognormal features

T	Specificity						
	KNN	RF	DT	SVM	LR	GB	XGB
1	61.1 (4.5)	63.2 (4.8)	57.1 (5.9)	**69.5** (3.5)	60.3 (2.7)	60.4 (4.3)	59.9 (5.0)
2	**58.0** (3.1)	54.0 (3.0)	57.4 (4.2)	55.6 (3.4)	57.3 (2.4)	54.3 (4.7)	54.8 (4.8)
3	61.9 (3.6)	77.8 (2.0)	67.6 (5.5)	**82.0** (2.6)	78.4 (2.0)	73.0 (3.2)	73.8 (3.3)
4	60.7 (4.0)	65.0 (3.9)	56.9 (6.4)	**68.5** (2.5)	66.5 (2.4)	62.8 (5.1)	63.8 (5.0)
5	71.3 (2.6)	73.4 (2.7)	64.8 (4.7)	**78.7** (2.9)	76.9 (3.1)	71.8 (3.5)	71.0 (3.9)
6	63.1 (3.6)	64.0 (4.0)	60.1 (5.0)	**69.8** (4.7)	67.1 (3.7)	65.4 (4.7)	67.3 (3.7)

Table 5. Average Precision achieved on 30 runs for every ML algorithm on lognormal features

T	Precision						
	KNN	RF	DT	SVM	LR	GB	XGB
1	65.3 (3.0)	65.7 (2.8)	59.8 (3.4)	**69.7** (2.7)	64.9 (1.7)	63.2 (2.7)	62.9 (3.1)
2	**66.3** (2.6)	66.2 (2.4)	65.0 (3.3)	63.8 (2.0)	65.2 (2.0)	65.9 (3.3)	65.4 (3.0)
3	63.2 (2.3)	76.4 (1.9)	68.5 (3.3)	**80.3** (2.5)	77.6 (2.0)	72.1 (2.8)	73.6 (3.0)
4	69.2 (2.6)	74.4 (2.5)	66.7 (3.7)	**74.5** (2.0)	74.3 (1.7)	72.8 (3.5)	73.2 (2.8)
5	69.6 (2.7)	73.8 (2.3)	64.8 (3.8)	**77.3** (2.8)	76.0 (2.9)	71.6 (3.6)	71.1 (3.0)
6	69.0 (3.1)	69.6 (3.1)	64.3 (3.9)	**74.6** (3.1)	72.8 (2.6)	71.1 (2.9)	72.3 (2.7)

Table 6. Average FNR achieved on 30 runs for every ML algorithm on lognormal features

T	FNR						
	KNN	RF	DT	SVM	LR	GB	XGB
1	32.7 (2.7)	35.6 (3.0)	40.7 (4.2)	35.4 (2.7)	**32.6** (2.1)	37.3 (2.5)	37.4 (3.1)
2	33.3 (3.5)	**27.9** (3.9)	36.4 (4.2)	37.4 (3.5)	35.4 (3.2)	29.4 (4.3)	32.0 (3.6)
3	36.4 (2.0)	31.9 (3.7)	33.6 (4.1)	32.6 (1.8)	**29.8** (2.0)	33.6 (4.2)	31.3 (3.5)
4	32.2 (3.0)	**22.5** (2.1)	33.8 (5.4)	31.1 (3.0)	26.5 (2.7)	24.4 (3.7)	24.0 (3.1)
5	37.5 (3.1)	**29.8** (3.0)	37.6 (4.1)	31.3 (2.1)	31.4 (2.1)	32.3 (4.2)	31.7 (4.3)
6	29.4 (3.5)	30.3 (4.6)	37.9 (6.2)	24.8 (3.4)	**24.7** (3.6)	28.8 (4.8)	27.4 (5.3)

Table 7. Average AUC achieved on 30 runs for every ML algorithm on lognormal features

T	AUC						
	KNN	RF	DT	SVM	LR	GB	XGB
1	0.66 (0.02)	0.70 (0.01)	0.58 (0.03)	**0.72** (0.02)	0.71 (0.01)	0.66 (0.02)	0.67 (0.02)
2	0.65 (0.02)	**0.68** (0.02)	0.60 (0.03)	0.63 (0.04)	0.65 (0.02)	0.67 (0.03)	0.66 (0.03)
3	0.69 (0.01)	0.82 (0.01)	0.66 (0.02)	0.82 (0.01)	**0.83** (0.01)	0.79 (0.02)	0.78 (0.02)
4	0.68 (0.02)	0.78 (0.01)	0.61 (0.03)	0.78 (0.01)	**0.79** (0.01)	0.76 (0.02)	0.77 (0.02)
5	0.69 (0.01)	**0.78** (0.01)	0.63 (0.02)	0.77 (0.01)	0.76 (0.01)	0.75 (0.02)	0.75 (0.02)
6	0.68 (0.02)	0.72 (0.03)	0.74 (0.02)	**0.76** (0.02)	0.75 (0.02)	0.73 (0.03)	0.74 (0.02)

in Sect. 3.1. The image generation process and the comparison between the our approach and DL are detailed in the following.

RGB Images. Starting from the raw data acquired as described in Sect. 3.1 and stored in terms of x-y coordinates and pressure at a frequency of 200 Hz, we generated synthetic images to feed Convolutional Neural Networks (CNN). The traits of these images are obtained by considering the points (x_i, y_i) as vertices of the polygonal that approximates the original curve. We encoded kinematic information in the RGB channels and as the tools used for the acquisition step allow us to record in air movements too, these images contains both in air and on paper information. In particular, they were obtained by considering the triplet of values (z_i, v_i, j_i) assumed as RGB color components for the $i-th$ trait, delimited by the couple of points (x_i, y_i) and (x_{i+1}, y_{i+1}). The triplet is obtained as follows:

- z_i is the pressure value at point (x_i, y_i) and it is assumed constant along the i-th trait;
- v_i is the velocity of the i-th trait, computed as the ratio between the length of the i-th trait and interval time of 5 ms corresponding to the period of acquisition of the tablet;
- j_i is the jerk of the i-th trait, defined as the second derivative of velocity.

The values of the triplet (z_i, v_i, j_i) have been normalized into the range [0, 255] in order to match the standard 0–255 color scale, by considering the minimum and the maximum value on the entire training set for these three quantities. For further details about the generation of these images, we suggest checking out our recent publication [3]. We selected three CNN models that accept input images that are automatically resized to 256×256 for VGG19 [22], to 224×224 for ResNet50 [7], to 299×299 for InceptionV3 [23] respectively. Taking into account these constraints for both type of images, the original x, y coordinates have been resized into the range [0, 299] for each image, in order to provide ex-ante images of suitable size and minimize the loss of information related to possible zoom in/out.

ML/DL Comparison. As mentioned above, lognormal features can be given in input to standard ML algorithms, whereas RGB images contain dynamic information encoded into the three color channels and can be used to feed a different CNN. Table 8 shows the accuracy performances achieved by the two approaches. From the table we can observe that in most cases ML outperformed DL, especially with the SVM classifier. DL only won on the second task with the VGG19 net. For the sake of comparison, for each task we plotted the ROC curves of the classification algorithms/nets that outperformed the others in at least one task, namely LR and SVM among the ML classifiers, and VGG19 among the CNNs (see Table 8). Looking at these two different sources of evaluation, we can observe that the deep approach (RGB images) outperformed the lognormal-based one on the graphic tasks (Tasks #1 and #2). On the contrary, the lognormal features confirmed their effectiveness in dealing with handwriting and cognitive tasks (see Fig. 2).
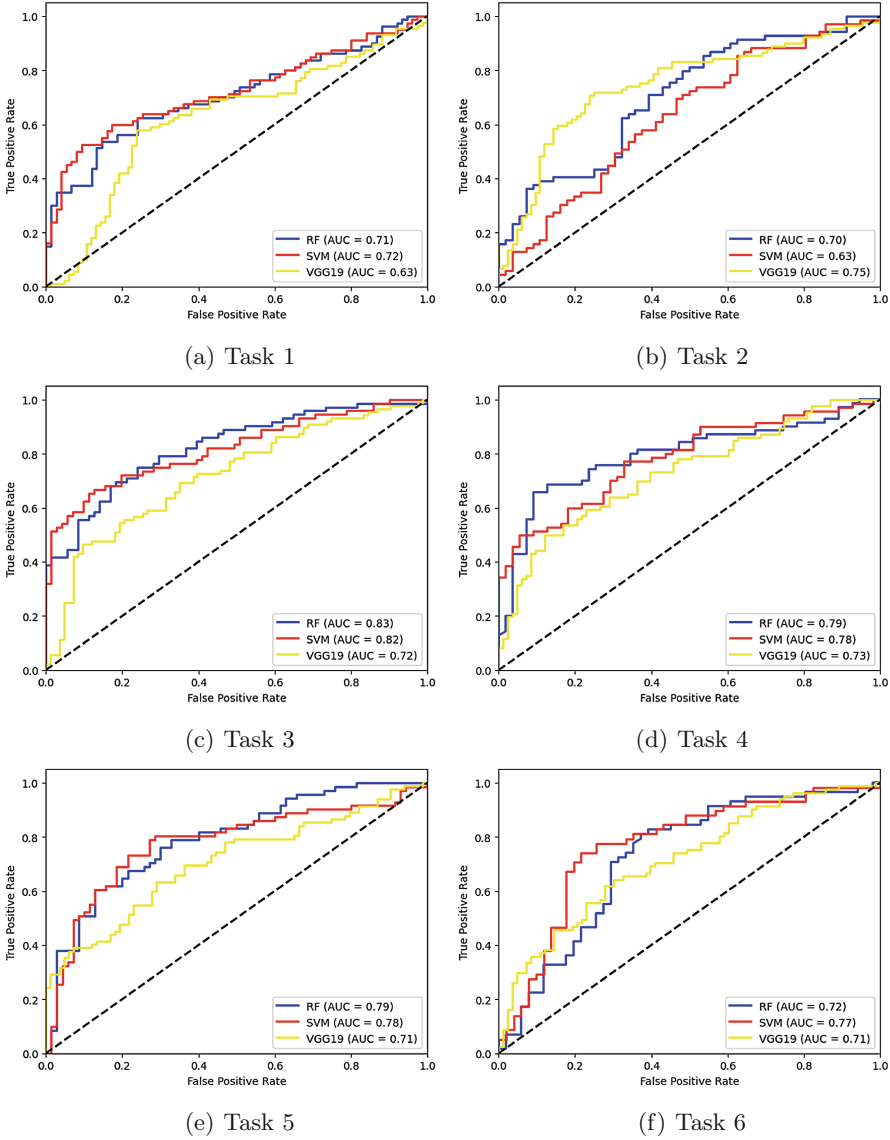
Fig. 2. Comparison of ROC curves obtained from RF, SVM and VGG19 for every task.

Table 8. Comparison results.

T	ML							Deep		
	KNN	RF	DT	SVM	LR	GB	XGB	VGG19	ResNet50	Inc.V3
1	64.34	63.83	58.24	**66.92**	63.98	61.56	61.33	61.62	62.64	62.20
2	62.77	63.97	60.83	59.47	61.31	63.30	62.05	**72.19**	65.90	71.25
3	62.79	72.97	67.01	**74.66**	74.25	69.73	71.27	66.83	62.09	70.81
4	64.67	**72.03**	62.11	68.72	70.43	69.97	70.59	66.82	58.62	63.97
5	66.87	71.80	63.58	**73.69**	72.68	69.70	69.63	66.01	62.43	70.37
6	67.09	67.10	61.14	**72.66**	71.50	68.45	70.10	66.48	64.07	65.39

5   Conclusions and Future Work

Neurodegenerative disease is a cognitive impairment that can be manifested through the graphonomics lack of skills. Alzheimer's and Parkinson's are the two most common diseases observed in handwriting. This paper analyzes the writing of patients and healthy people with kinematic features extracted from the kinematic theory of rapid movement. To study the skill levels of participants, we used a dataset with healthy people and patients at the early stage of AD. Their handwriting included signatures, letters, and drawings following a well-established protocol for Alzheimer's [4]. Our preliminary results confirm that lognormal features model handwriting better than graphic tasks. In particular, we achieved the best results (ACC > 70 %) with the following tasks:

– 'le' (repeated four times);
– Word 'bottiglia' backward;
– Dictated telephone number;
– Clock drawing test.

It is worth noting that we did not apply any parameter optimization in this work. We expect that a grid search procedure will allow us a significant improvement in our future outcomes. Furthermore, these results align with those achieved using standard kinematic features (velocity, acceleration, etc.). We also plan to analyze which tasks lognormal/deep features perform better. This would allow us deepening for each task which features (approach) achieve the best classification performance. Finally, similar to related works, there is room to improve our final prediction by combining the responses from the single classifiers (one per task).

References

1. Albu, A., Precup, R., Teban, T.A.: Results and challenges of artificial neural networks used for decision making and control in medical applications. Facta Univ. Ser. Mech. Eng. **17**(3), 285–308 (2019)

2. Carmona-Duarte, C., Ferrer, M., Plamondon, R., Gómez-Rodellar, A., Gómez-Vilda, P.: Sigma-lognormal modeling of speech. Cogn. Comput. **13**(2), 488–503 (2021). https://doi.org/10.1007/s12559-020-09803-8

3. Cilia, N., D'Alessandro, T., De Stefano, C., Fontanella, F., Molinara, M.: From online handwriting to synthetic images for Alzheimer's disease detection using a deep transfer learning approach. IEEE J. Biomed. Health Inform. **25**(12), 4243–4254 (2021). https://doi.org/10.1109/JBHI.2021.3101982

4. Cilia, N.D., De Stefano, C., Fontanella, F., Di Freca, A.S.: An experimental protocol to support cognitive impairment diagnosis by using handwriting analysis. Proc. Comput. Sci. **141**, 466–471 (2018)

5. Djioua, M., Plamondon, R.: A new algorithm and system for the characterization of handwriting strokes with delta-lognormal parameters. IEEE Trans. Pattern Anal. Mach. Intell. **31**(11), 2060–2072 (2009)

6. Ferrer, M.A., Diaz, M., Carmona-Duarte, C., Plamondon, R.: IDeLog: iterative dual spatial and kinematic extraction of sigma-lognormal parameters. IEEE Trans. Pattern Anal. Mach. Intell. **42**(1), 114–125 (2020)

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

8. Impedovo, D.: Velocity-based signal features for the assessment of parkinsonian handwriting. IEEE Signal Process. Lett. **26**(4), 632–636 (2019)

9. Impedovo, D., Pirlo, G., Balducci, F., Dentamaro, V., Sarcinella, L., Vessio, G.: Investigating the sigma-lognormal model for disease classification by handwriting. In: The Lognormality Principle and its Applications in E-Security, E-Learning and E-Health, pp. 195–209. World Scientific (2021)

10. Impedovo, D., Pirlo, G., Vessio, G.: Dynamic handwriting analysis for supporting earlier Parkinson's disease diagnosis. Information **9**(10), 247 (2018)

11. Lebel, K., Nguyen, H., Duval, C., Plamondon, R., Boissy, P.: Capturing the craniocaudal signature of a turn with inertial measurement systems: methods, parameters robustness and reliability. Front. Bioeng. Biotechnol. **5**, 51 (2017). https://doi.org/10.3389/fbioe.2017.00051

12. Myszczynska, M.A., et al.: Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. Nat. Rev. Neurol. **16**, 440–456 (2020)

13. O'Reilly, C., Plamondon, R.: Design of a neuromuscular disorders diagnostic system using human movement analysis. In: 2012 11th International Conference on Information Science, Signal Processing and their Applications, ISSPA 2012, pp. 787–792 (2012). https://doi.org/10.1109/ISSPA.2012.6310660. Cited by: 28

14. O'Reilly, C., Plamondon, R.: Development of a sigma-lognormal representation for on-line signatures. Pattern Recogn. **42**(12), 3324–3337 (2009). https://doi.org/10.1016/j.patcog.2008.10.017. new Frontiers in Handwriting Recognition

15. Plamondon, R.: A kinematic theory of rapid human movements - part II. Movement time and control. Biol. Cybern. **72**(4), 309–320 (1995). https://doi.org/10.1007/BF00202786

16. Plamondon, R.: A kinematic theory of rapid human movements: part I. Movement representation and generation. Biol. Cybern. **72**(4), 295–307 (1995)

17. Plamondon, R.: A kinematic theory of rapid human movements: part III. Kinetic outcomes. Biol. Cybern. **78**(2), 133–145 (1998). https://doi.org/10.1007/s004220050420

18. Plamondon, R., O'Reilly, C., Rémi, C., Duval, T.: The lognormal handwriter: learning, performing, and declining. Front. Psychol. **4** (2013). https://doi.org/10.3389/fpsyg.2013.00945, https://www.frontiersin.org/article/10.3389/fpsyg.2013.00945

19. Plamondon, R., Pirlo, G., Anquetil, É., Rémi, C., Teulings, H.L., Nakagawa, M.: Personal digital bodyguards for e-security, e-learning and e-health: a prospective survey. Pattern Recogn. **81**, 633–659 (2018). https://doi.org/10.1016/j.patcog.2018.04.012

20. Precup, R.E., Teban, T.A., Albu, A., Borlea, A.B., Zamfirache, I.A., Petriu, E.M.: Evolving fuzzy models for prosthetic hand myoelectric-based control. IEEE Trans. Instrum. Meas. **69**(7), 4625–4636 (2020)

21. Schröter, A., Mergl, R., Bürger, K., Hampel, H., Möller, H.J., Hegerl, U.: Kinematic analysis of handwriting movements in patients with Alzheimer's disease, mild cognitive impairment, depression and healthy subjects. Dement. Geriatr. Cogn. Disord. **15**(3), 132–42 (2003)

22. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. CoRR arXiv:1409.1556 (2015)

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

24. Tanveer, M., et al.: Machine learning techniques for the diagnosis of Alzheimer's disease: a review. ACM Trans. Multimed. Comput. Commun. Appl. **16**(1s), 1–35 (2020)

25. Tseng, M.H., Cermak, S.A.: The influence of ergonomic factors and perceptual-motor abilities on handwriting performance. Am. J. Occup. Ther. **47**(10), 919–926 (1993)

26. Vyhnálek, M., et al.: Clock drawing test in screening for Alzheimer's dementia and mild cognitive impairment in clinical practice. Int. J. Geriatr. Psychiatry **32**(9), 933–939 (2017)

27. Yan, J.H., Rountree, S., Massman, P., Smith Doody, R., Li, H.: Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. J. Psychiatr. Res. **42**(14), 1203–1212 (2008)