



Using Genetic Algorithms for the Prediction of Cognitive Impairments

Nicole Dalia Cilia, Claudio De Stefano, Francesco Fontanella^(✉),
and Alessandra Scotto Di Freca

Department of Electrical and Information Engineering (DIEI),
Università di Cassino e del Lazio Meridionale,
Via G. Di Biasio, 43, 03043 Cassino, FR, Italy
{nicoledalia.cilia,destefano,fontanella,a.scotto}@unicas.it

Abstract. Cognitive impairments affect millions of persons worldwide, and especially elderly ones. These impairments may be one of the first signs of the arising of neurodegenerative diseases, such as Alzheimer's and Parkinson's, and it is expected that the incidence of this kind of diseases will dramatically increase worldwide in the near future. For this reason, the improvement of the tools currently used to diagnose these diseases is becoming crucial. Handwriting is one of the human skills affected by this kind of impairments, and anomalies such as micrographia have been adopted as diagnosis sign for the Parkinson's disease. In a previous paper, we presented a study in which the handwriting of the subjects involved was recorded while they were performing some elementary tasks, such as the copy of simple words or the drawing of elementary forms. Then we extracted the features characterizing the dynamics of the handwriting and used them to train a classifier to predict whether the subject analyzed was affected by a cognitive impairment or not. In this paper, we present a system that uses a genetic algorithm to improve of the performance of the system previously presented. The genetic algorithm has been used to select the subset of tasks that allow improving the prediction ability of the previous system. The experimental results confirmed the effectiveness of the proposed approach.

Keywords: Genetic algorithms · Handwriting · Alzheimer's disease · Parkinson's disease · Neurodegenerative disorders

1 Introduction

Cognitive impairments are defined as cognitive decline greater than expected for an individual's age and education level but that does not interfere deeply with activities of daily life. Their symptoms can remain stable or even disappear, but for more than half of the cases they evolve into dementia diseases [10], thus their early identification could lead to the prevention of dementia diseases. Moreover, some types of cognitive impairments have a high risk of progression to Alzheimer's disease (AD), and then they can be considered as prodromal

symptoms of this disorder. The risk of being affected by Alzheimer's increases strongly with age thus it is expected that in the next decades the incidence of cognitive impairments will dramatically increase [19]. To date, Alzheimer's clinical diagnosis is performed by physicians that, in most cases, in addition to cognitive tests, perform some invasive biomarker tests, e.g. cerebrospinal fluid tests. This creates a strong need for the improvement of the approaches currently being used for diagnosis of these diseases.

One of the basic skills compromised by cognitive impairments is certainly the handwriting [3, 16, 17, 24]. In fact, it has been observed that some handwriting anomalies can be used as diagnostic signs. For example, has been found that the handwriting of Alzheimer's patients shows alterations in spatial organization and poor control of movement [18, 20]. Most of the studies which analyze the effects of cognitive impairments on handwriting have been conducted in the medical field, with few studies adopting classification algorithms to analyze people's handwriting to detect those affected by cognitive impairments. Moreover, almost all of these studies have involved a few dozens of subjects, thus limiting the effectiveness of classification algorithms. To try to overcome these problems, in [8] the authors proposed a protocol consisting of twenty five handwriting tasks, with the aim of investigating how cognitive impairments affect the different motor and cognitive skills involved in the handwriting process. The protocol has been then adopted to collect handwriting data from about one hundred seventy-five subjects, both cognitive impaired or healthy. This data has been then used to build a system in which two kinds of pen traits were distinguished, i.e. on-paper and on-air. With the former representing the movements in which the pen is touching the paper, whereas the latter are those in which the pen is on air during the handwriting process. It is worth noting that in the literature has been found that this distinction allows a better characterization of the handwriting anomalies caused by cognitive impairments [9, 13, 14]. For each task, the features extracted have been employed to train two classifiers, both for on-paper and on-air features [6, 7]. Finally, in order to predict the cognitive state of a subject (healthy or impaired) the fifty responses provided by the classifiers trained on the single tasks were combined according to the Majority-vote rule.

In the scenario outlined above, it arises the need to optimize system performance. The need is twofold: from one hand we want to maximize the prediction performance of the system; on the other hand, we want to reduce the number of tasks to be performed. This optimization problem can be seen as a combinatorial one, in which given the set $T = \{t_1, t_2, \dots, t_n\}$, the best subset must be found, according to a given evaluation function (the prediction performance in our case). Though the best subset can be found by exhaustively evaluating all the possible solutions, this search strategy is impracticable in our case, where the total number of possible solutions is $2^{50} \approx 10^{15}$. For this reason, a heuristic search is needed. Since Genetic Algorithms (GAs) are well-known for their global search ability without using any domain knowledge or assumptions about the search space, they have been widely used for this kind of combinatorial problems. This is because GA binary vectors provide a natural and straightforward

representation for item subsets: the value 1 or 0 of the chromosome i -th element indicates whether the i -th item is included or not in the subset represented by the chromosome.

Because of their search-ability Evolutionary algorithms have been also used for health applications. In particular, GA and GP have been mostly used. As concerns the GP-based approaches, they have been used in a wide range of applications. For example, in [2], the authors proposed a constrained-syntax GP-based algorithm for discovering classification rules in medical data. The authors tested their approach on five datasets and achieved better results than decision trees. In [4], instead, the authors tackled a problem related to the physio-chemical properties of proteins, involving the prediction on these properties in tertiary structure. The authors proved that the proposed approach was more effective than artificial Neural Networks and Support Vector Machines. More recently, the Cartesian GP approach has been used to automatically identify subjects affected by the Parkinson's disease through the analysis of their handwriting [22, 23]. GP algorithms have been also used as a tool to support medical decisions for treating rare diseases [1].

Also GA has been widely used in medical applications, in most of the medical specialities, e.g., medical imaging, rehabilitation medicine, and health care management [11]. As concerns neurodegenerative diseases, in [25] the authors used a GA with Multi-Objective fitness function to find the relevant volumes of the brain related to the Alzheimer's disease, whereas in [15] the authors used a GA to search the optimal set of neuropsychological tests, to be used to build a system for the prediction of the Alzheimer's disease.

Note that, to the best of our knowledge, there are no other studies in which a GA has been used to improve the performance of a system for the prediction of cognitive impairments, based on handwriting analysis.

In this paper, we present a GA-based system for the improvement of the prediction performance of the system described above. In particular, the devised system selects the subset of tasks, both on-air and on paper, that allows improving the performance of the proposed system in predicting the cognitive impairments of the subjects involved. The final prediction of the cognitive state of a subject is made applying the majority vote rule to the responses collected from the selected tasks. The system has been trained by using the dataset made of the responses provided by the classifier used to predict the cognitive impairment of a subject for a given task. A part of it has been used as a training set, to implement the fitness function for evaluating the individuals to be evolved, whereas the remaining part has been used as a test set to assess the performance of the system on unseen data.

We tested our approach taking into account four well-known classifiers: decision trees, random forests, neural networks, and support vector machines. Moreover, in the first set of experiments, we analyzed the generalization ability of the system as well as its capability in reducing the number of tasks needed to correctly predict cognitive impairments. Then, by counting the occurrences of the tasks in the solutions found in the several runs performed, we tried to figure out

the more relevant tasks. Finally, we compared the proposed approach with the majority-vote and the weighted majority rules, which are well-known and widely used strategies for combining the responses provided by several classifiers. Moreover, since the problem in hand is a combinatorial one in which the best subset must be found, as is the case of feature selection problems [5], we compared our results with those achieved by a state-of-the-art algorithm for feature selection. The comparison results confirmed the effectiveness of the proposed approach.

The remainder of the paper is organized as follows: Sect. 2 describes the data collection and the protocol developed to collect handwriting data. Section 3 details the proposed system. Section 4 displays the experiments and presents the results achieved. Finally, Sect. 5 is devoted to some concluding remarks.

2 Data Collection and Protocol

In the following subsections, the data collection procedure, the protocol designed for collecting handwriting samples, the segmentation and feature extraction methods, are detailed.

2.1 Data Collection

The one hundred seventy five subjects who participated in the experiments, namely eighty six cognitively impaired patients and eighty nine healthy controls, were recruited with the support of the geriatric ward, Alzheimer unit, of the University hospital Federico II in Naples. As recruiting criteria we have took into account clinical tests (such as PET, TAC and enzymatic analyses) and standard cognitive tests (such as MMSE). 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.

The data were collected by using a graphics tablet, which allowed the subjects to write on standard A4 white sheets using an apparently normal pen: such a pen produces both the ink trace on the sheet and the digital information, which are recorded on the tablet in the form of spatial coordinates and pressure for each point, acquired at a frequency of 200 Hz. The tablet also records the in-air movements (up to a maximum of three centimetres in height). In this condition, the subject should not change his natural writing movements as it happens, for instance, when the writing is produced with a stylus on the surface of a tablet.

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 cognitively impaired subjects from the healthy ones. The goal of these tasks is to test the patients' abilities in repeating simple as well as complex graphic gestures, which have a semantic meaning, such as letters and words of different lengths and with different spatial organizations. The tasks considered for this

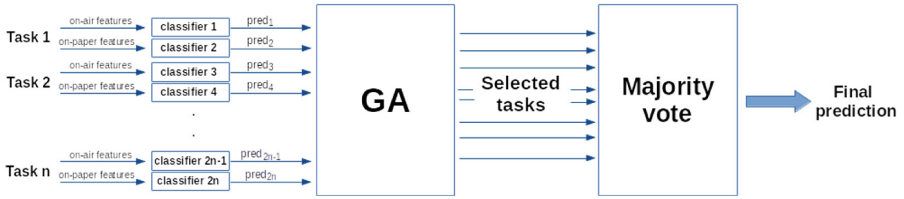


Fig. 1. The layout of the proposed system. Note that in our case $n = 25$.

study are been presented in [8], and they are arranged in increasing order of difficulty, in terms of the cognitive functions required. Taking into account their objectives, we have grouped the tasks as follows:

- *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).
- *Copy* and *Reverse Copy* tasks, whose objective is to test the patient’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).
- *Memory* tasks, whose objective is to test the variation of the graphic section, keeping in memory a word, a letter, a graphic gesture or a motor planning.
- *Dictation*, whose purpose is to investigate how the writing in the task varies (with phrases or numbers) in which the use of the working memory is necessary.

3 The Proposed System

The following subsections details the steps performed to implement the proposed system (see Fig. 1).

3.1 Segmentation and Feature Extraction

We extracted both on-paper and on-air features by considering their segmentation in elementary strokes assuming as segmentation points both pen-up and pen-down, as well as the zero-crossing of the vertical velocity profile. The feature values were computed for each stroke and averaged over all the strokes relative to a single task. In particular, from each task, we have extracted the following types of features:

- (i) Static features: Start time; Duration; Initial vertical position; Vertical dimension; Initial horizontal position; Horizontal dimension; Inclination from the initial point to the final point; Loop surface; Absolute size.

- (ii) Dynamic features: Vertical speed peak; Peak of vertical acceleration; Straightness error; Relative initial inclination; Time relative to the vertical speed peak; Relative duration of the on-paper sections; Average speed; Absolute Jerk: the Root Mean Square (RMS) value of the absolute jerk on all samples of a stroke or segment; Normalized Jerk; Number of points of acceleration peaks; Average pen pressure.

Note that we have separately redefined and computed the features over on-paper and on-air traits since in the literature has been found that the two conditions exhibit significant differences in characterizing the handwriting anomalies caused by cognitive impairments [9].

3.2 Classification

Once the features were extracted, for each task, two classifiers were trained (both for on-paper and on-air features), and the test predictions obtained using the 5-fold cross-validation was stored. As a consequence, at the end of this step, for each classification algorithm considered, a dataset containing one hundred seventy-five samples was available, one for each subject, with each sample consisting of fifty predictions. The datasets so built were used to train and test the GA module detailed in the next subsection. In particular, a part of it was used as the training set (T_r in the following), to implement the fitness function for evaluating the individuals to be evolved, whereas the remaining part was used as test set (T_s) to assess the performance of the system on unseen data.

3.3 GA for Task Selection

To optimize the performance of the system in predicting the cognitive state of the involved subjects, we used a GA to select the subset of fifty tasks (both on-air and on-paper) that allows the system to achieve the best prediction performance. As mentioned in the introduction, we used a GA because this algorithm is well-known for its global search ability and also because it provides a natural and straightforward representation of item (tasks in our case) subsets: the value 1 or 0 of the chromosome i -th element indicates whether the i -th item (task) is included or not in the subset represented by the chromosome. Given the i -th individual to be evaluated, representing the task subset s_i , its fitness was computed by considering in the majority-vote rule only the tasks included in s_i (see Fig. 1).

The GA was implemented by using a generational evolutionary algorithm, which starts by generating a population of P individuals, randomly generated. Afterwards, the fitness of the generated individuals is evaluated by computing the prediction accuracy on T_r . After this preliminary evaluation phase, a new population is generated by selecting $P/2$ couples of individuals using the tournament method, of size t . The one point crossover operator is then applied to each of the selected couples, according to a given probability factor p_c . Afterwards, the mutation operator is applied with a probability p_m . The value of p_m was

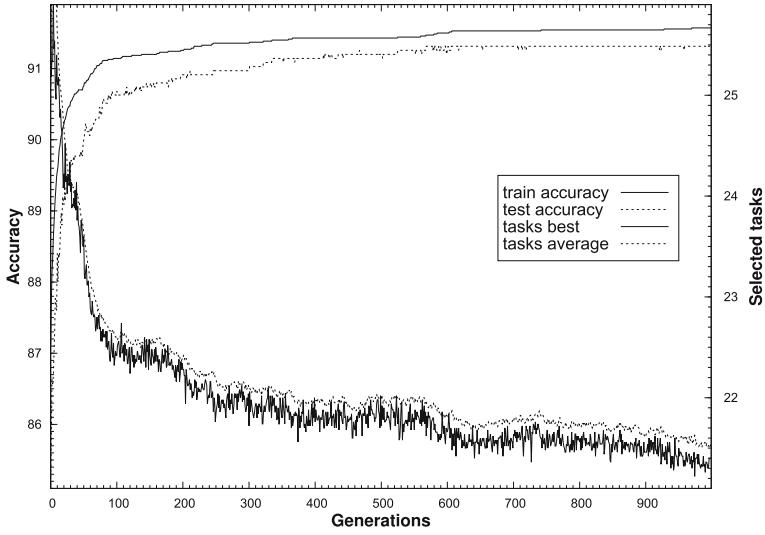
set to $1/50$, where 50 is the chromosome length, i.e. the total number of available tasks. This probability value allows, on average, the modification of only one chromosome element. This value has been suggested in [21] as the optimal mutation rate below the error threshold of replication. Finally, these individuals are added to the new population. The process just described is repeated for N_g generations.

4 Experimental Results

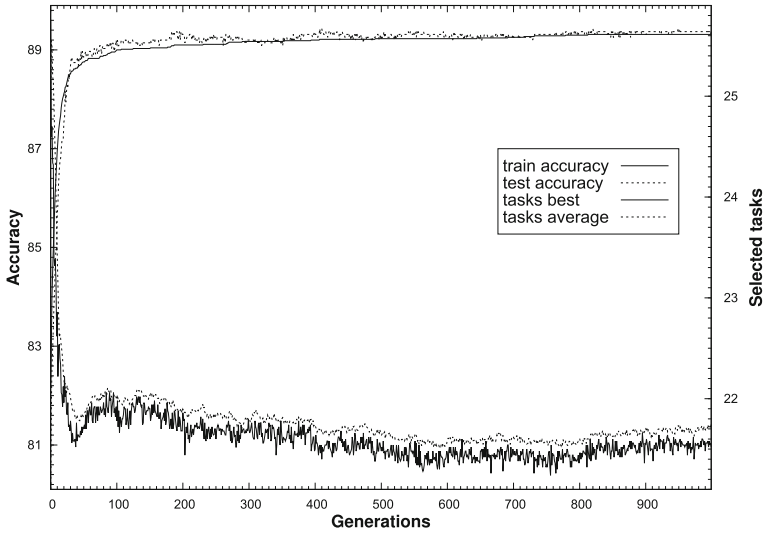
We tested our approach by training four well-known and widely-used classifiers: decision trees (DT), random forests (RF), neural networks (NN), and support vector machines (SVM). Thus, according to the procedure detailed in Subsect. 3.2, we built up four datasets, where each sample contained the responses provided by the given classifier on the fifty tasks, i.e. twenty five for the on-paper features and twenty five for the on-air features (see Fig. 1). Each dataset was split into two parts, statistically independent: a training set T_r , made of the 80% of the available samples, and a test set T_s , made of the remaining samples. T_r was used to evaluate the individuals' fitness, whereas T_s was used to assess the performance of the best individual on unseen data. For each dataset, we performed fifty runs and at the end of each run, the task subset encoded by the individual with the best fitness was stored as the solution provided by that run. The results reported in the following were computed by averaging those obtained by the fifty best individuals stored. As for the parameters of the GA, we performed some preliminary trials to set them. These parameters were used for all the experiments described below and are reported in Table 1. We performed three sets of experiments. In the first set, we analyzed the generalization ability of the GA-based system as well as its capability in reducing the number of tasks needed to correctly predict cognitive impairments. To this aim, we have plotted the training and test accuracy of the best individual as well as the population's average number of selected tasks and the number of the selected tasks of the best individual during the fifty runs performed. As concerns the second set of experiments, we analyzed the number of occurrences of the selected tasks in order to figure out which are the most relevant ones. Finally, we compared the results

Table 1. The values of the parameters used in the experiments.

Parameter	Symbol	Value
Population size	P	100
Crossover probability	p_c	0.6
Tournament size	t	5
Elitism	e	2
Mutation probability	p_m	0.02
Number of Generations	N_g	1000



(a) DT

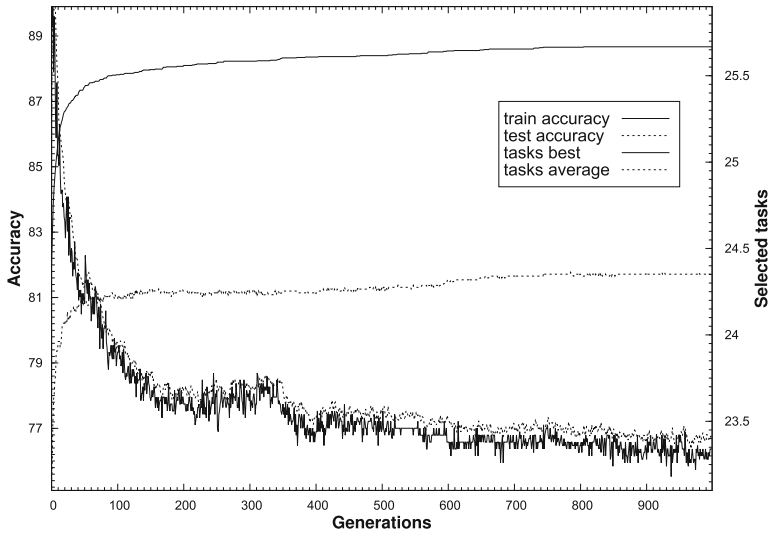


(b) NN

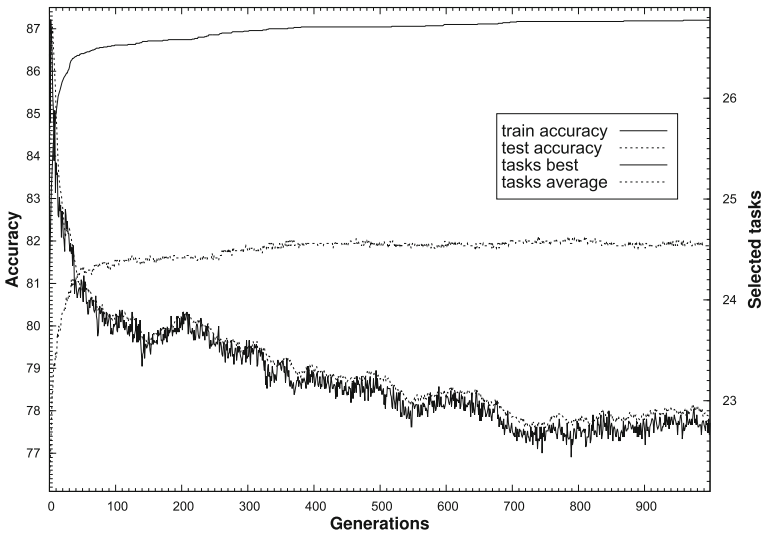
Fig. 2. Evolution of accuracy and average number of selected tasks for DT and NN classifiers.

achieved by the proposed approach with those obtained by the majority-vote and weighted majority rules as well as those achieved by the floating Forward Selection (FFS) algorithm.

The plots obtained from the first set of experiments are shown in Figs. 2 and 3. The plots show the evolution of: (i) the average training and test accuracy



(a) RF



(b) SVM

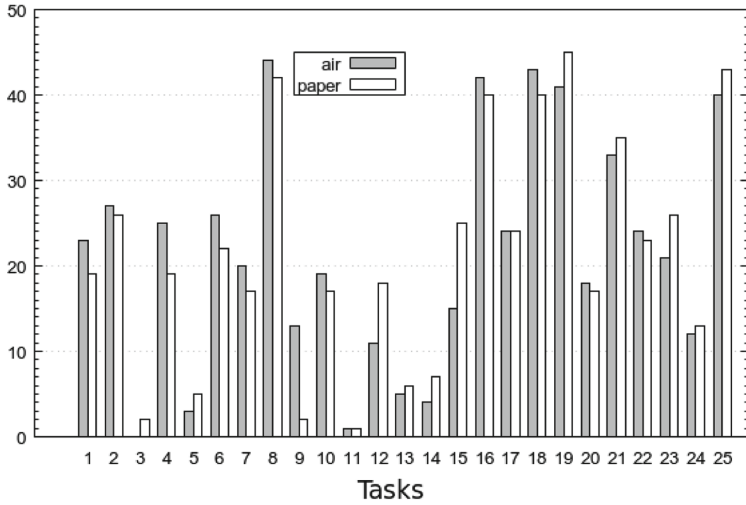
Fig. 3. Evolution of accuracy and average number of selected tasks for RF and SVM classifiers.

of the best individual; (ii) the average number of selected tasks for the best individual and the whole population, computed by averaging the values of the fifty runs performed. From the plots, it can be seen that, for every classifier, the GA did not suffer overfitting, because the test and train curves show similar trends, although train and test accuracies differ a lot among the classifiers. Most

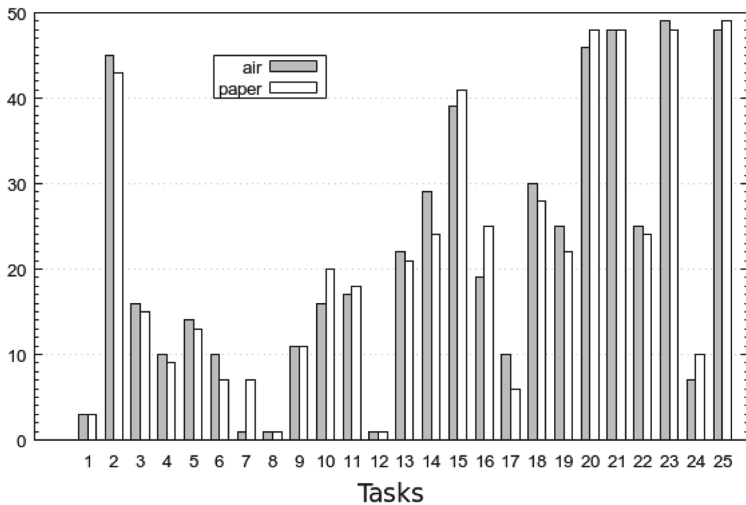
probably, these differences depend on the generalization ability of the classifiers, with the DT and NN exhibiting very good results, whereas RF and SVM achieved poor generalization results. A common, and very interesting, trend emerging from the plots is that as the number of iterations increases the fitness function increases, even if only slightly, and the number of selected tasks, both average and best, decreases. At the end of the run, for every classifier, about less than half of the tasks were selected. Moreover, it is worth noting that the number of selected tasks of the best individual is always less than the average one. This seems to confirm our assumption that better prediction performance can be achieved by suitably selecting a subset of the whole set of available tasks.

The histograms of the second set of experiments are shown in Figs. 4 and 5. They show the number of occurrences of the selected tasks, computed on the fifty runs performed. Gray and white bars represent on-air and on-paper tasks, respectively. The histograms does not show a common pattern, with some tasks more selected for a given classifier and less or even far less for the remaining ones, as is the case, for example, for the task 1 (signature), which was selected about twenty times for the DT and far less for the remaining classifiers. This depends on the fact that for a given task the classifiers achieved different prediction performances. The only aspect shared by the four histograms is that for the same task the number of occurrences of the on-paper and on-air tasks differs little. This is due to the fact that, for every task, for a given classifier, the prediction performance of on-air and on-paper features are very similar.

In order to test the effectiveness of our system, we compared its results with those achieved by the majority-vote and the weighted majority rules, which are well-known and widely used strategies for combining the responses provided by several classifiers. The first rule, given a set of responses to be combined, in our case the prediction provided by the classifiers for the single tasks, assigns an unknown sample (a subject in our case) to the class (CI or HC) that has the highest occurrence among those provided by the whole set of classifiers. As for the second rule, it takes into account the prediction performance of the single classifiers by multiplying each response with the training accuracy achieved by the related classifier. Finally, since the problem in hand is a combinatorial one in which the best subset must be found, as is the case of feature selection problems, we compared our results with those achieved by the Sequential Forward Floating Search Algorithm (SFS in the following). This strategy searches the solution space by using a greedy hill-climbing technique. It starts with the empty set of features and, at each step, selects the best feature according to the subset evaluation function. SFS stops when the addition of a new feature does not produce any improvement. Further details can be found in [12]. To statistically validate the comparison results, we performed the non-parametric Wilcoxon rank-sum test ($\alpha = 0.05$). Comparison results are shown in Table 2, where the values in bold are the best ones, for a given classifier, according to the Wilcoxon test. From the table, it can be seen that the DT achieved the best overall prediction performance. Moreover, the GA achieved much better results than those of the combination rules. The GA largely outperforms SFS, on all the four classifiers



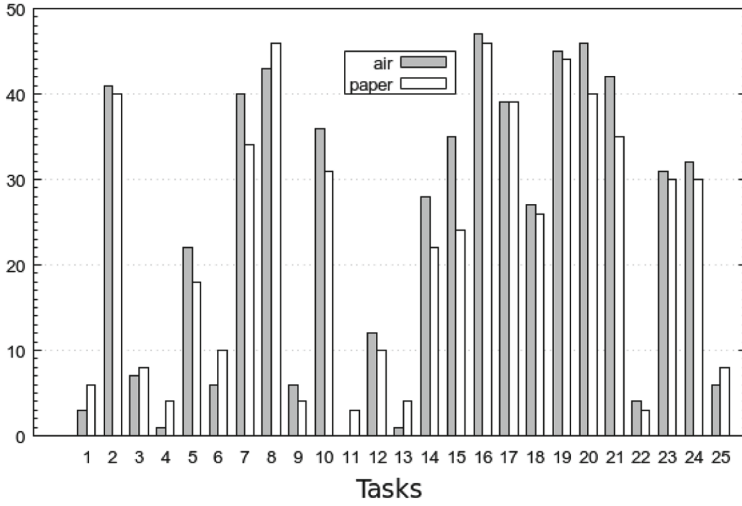
(a) DT



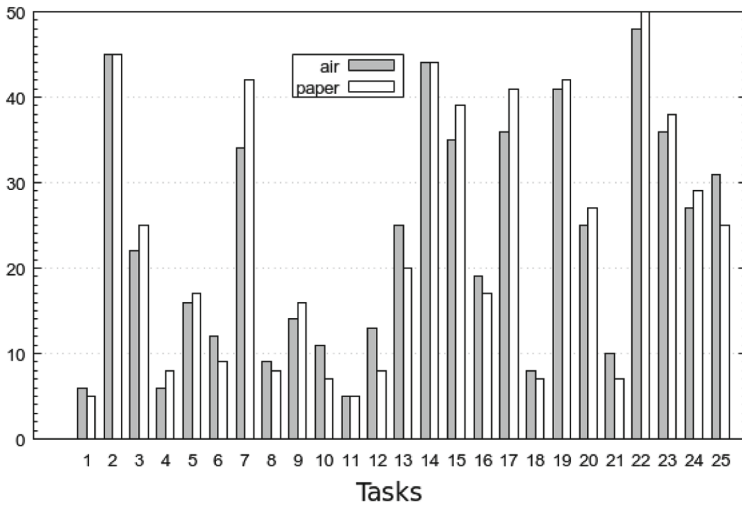
(b) NN

Fig. 4. Number of occurrences of the selected tasks for DT and NN classifiers.

considered, confirming its effectiveness. Finally, it is worth noting that the standard deviations exhibited by the GA results are significantly lower than those of the other results confirming the average good quality of solutions found by our system.



(a) RF



(b) SVM

Fig. 5. Number of occurrences of the selected tasks for RF and SVM classifiers.

Table 2. Comparison results. For each classifier the best result is highlighted in bold.

Classifier	GA		MjV		WMjV		FFS	
	Avg	Dev	Avg	Dev	Avg	Dev	Avg	Dev
DT	91.26	1.75	83.40	5.89	83.43	5.8	68.86	7.1
NN	89.20	2.23	72.00	6.93	73.34	6.3	77.02	3.6
RF	81.71	2.29	67.83	5.68	68.13	5.68	76.69	2.0
SVM	82.11	2.05	79.09	5.98	79.14	5.92	70.05	7.3

5 Conclusions

Cognitive impairments are one the first signs of the arising of neurodegenerative diseases and it is expected their incidence will increase in the near future. Thus, the improvement of the tools currently available to diagnose these diseases is becoming crucial. Since handwriting is one of the human skills affected by cognitive impairments, it can be analyzed to detect this kind of diseases.

In this paper, we presented a GA-based approach for the improvement of the performance of a system that record and analyze the handwriting of the involved subjects in performing some simple tasks, in order to detect those that are cognitively impaired. In particular, the system selects the subset of tasks that allow the maximization of the prediction performance.

In the experiments performed we tested the generalization ability of the system as well as its capability in reducing the number of tasks needed to correctly predict cognitive impairments. Moreover, we also tried to figure out which are the more relevant tasks, i.e. those most frequently selected by the GA in the performed runs. Finally, we compared our results with those of the majority-vote and the weighted majority rules, as well as the sequential floating search algorithm. The experimental results showed that the proposed system has a good generalization ability, by selecting only half of the available tasks. As concerns the relevance of the single tasks, the results showed that it varies greatly among the different classifiers considered; most probably, this depends on the fact that, for a given task, they exhibit a wide performance variability. As concerns the comparison results, they confirmed the effectiveness of our system.

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

References

1. Bakurov, I., Castelli, M., Vanneschi, L., Freitas, M.J.: Supporting medical decisions for treating rare diseases through genetic programming. In: Kaufmann, P., Castillo, P.A. (eds.) *EvoApplications 2019*. LNCS, vol. 11454, pp. 187–203. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16692-2_13

2. Bojarczuk, C.C., Lopes, H.S., Freitas, A.A., Michalkiewicz, E.L.: A constrained-syntax genetic programming system for discovering classification rules: application to medical data sets. *Artif. Intell. Med.* **30**(1), 27–48 (2004)
3. Carmona-Duarte, C., Ferrer, M.A., Parziale, A., Marcelli, A.: Temporal evolution in synthetic handwriting. *Pattern Recogn.* **68**(Suppl. C), 233–244 (2017)
4. Castelli, M., Vanneschi, L., Manzoni, L., Popovič, A.: Semantic genetic programming for fast and accurate data knowledge discovery. *Swarm Evol. Comput.* **26**, 1–7 (2016)
5. Cilia, N., De Stefano, C., Fontanella, F., di Freca, A.S.: A ranking-based feature selection approach for handwritten character recognition. *Pattern Recognit. Lett.* (2018)
6. Cilia, N.D., De Stefano, C., Fontanella, F., Molinara, M., Di Freca, A.S.: Handwriting analysis to support alzheimer's disease diagnosis: a preliminary study. In: Vento, M., Percannella, G. (eds.) CAIP 2019. LNCS, vol. 11679, pp. 143–151. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29891-3_13
7. Cilia, N.D., De Stefano, C., Fontanella, F., Molinara, M., Di Freca, A.S.: Using handwriting features to characterize cognitive impairment. In: Ricci, E., Rota Bulò, S., Snoek, C., Lanz, O., Messelodi, S., Sebe, N. (eds.) ICIAP 2019. LNCS, vol. 11752, pp. 683–693. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30645-8_62
8. Cilia, N., De Stefano, C., Fontanella, F., Di Freca, A.S.: An experimental protocol to support cognitive impairment diagnosis by using handwriting analysis. In: *Proceeding of The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH)*, pp. 1–9. Elsevier (2019). *Procedia Comput. Sci*
9. De Stefano, C., Fontanella, F., Impedovo, D., Pirlo, G., di Freca, A.S.: Handwriting analysis to support neurodegenerative diseases diagnosis: a review. *Pattern Recognit. Lett.* **121**, 37–45 (2019)
10. Elbaz, A., Carcaillon, L., Kab, S., Moisan, F.: Epidemiology of Parkinson's disease. *Revue Neurologique* **172**(1), 14–26 (2016)
11. Ghaheri, A., Shoar, S., Naderan, M., Hoseini, S.S.: The applications of genetic algorithms in medicine. *Oman Med. J.* **30**(6), 406–416 (2015)
12. Gütlein, M., Frank, E., Hall, M., Karwath, A.: Large scale attribute selection using wrappers. In: *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2009)* (2009)
13. Impedovo, D., Pirlo, G.: Dynamic handwriting analysis for the assessment of neurodegenerative diseases: a pattern recognition perspective. *IEEE Rev. Biomed. Eng.*, 1–13 (2018)
14. Impedovo, D., Pirlo, G., Vessio, G., Angelillo, M.T.: A handwriting-based protocol for assessing neurodegenerative dementia. *Cogn. Comput.* **11**(4), 576–586 (2019). <https://doi.org/10.1007/s12559-019-09642-2>
15. Johnson, P., et al.: Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease. *BMC Bioinform.* **15**(S11) (2014). <https://doi.org/10.1186/1471-2105-15-S16-S11>
16. Marcelli, A., Parziale, A., Santoro, A.: Modeling handwriting style: a preliminary investigation. In: *2012 International Conference on Frontiers in Handwriting Recognition*, pp. 411–416, September 2012
17. Marcelli, A., Parziale, A., Santoro, A.: Modelling visual appearance of handwriting. In: Petrosino, A. (ed.) ICIAP 2013. LNCS, vol. 8157, pp. 673–682. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-41184-7_68

18. Marcelli, A., Parziale, A., Senatore, R.: Some observations on handwriting from a motor learning perspective. In: 2nd International Workshop on Automated Forensic Handwriting Analysis (2013)
19. James, M., Wimo, A., Guerchet, M., Ali, G., Wu, Y.T., Prina, M.: World Alzheimer report 2015-the global impact of dementia: an analysis of prevalence, incidence, cost and trends. Alzheimer's Disease International, August 2015
20. Neils-Strunjas, J., Groves-Wright, K., Mashima, P., Harnish, S.: Dysgraphia in Alzheimer's disease: a review for clinical and research purposes. *J. Speech Lang. Hear. Res. Res.* **49**(6), 1313–30 (2006)
21. Ochoa, G.: Error thresholds in genetic algorithms. *Evol. Comput.* **14**(2), 157–182 (2006)
22. Senatore, R., Della Cioppa, A., Marcelli, A.: Automatic diagnosis of neurodegenerative diseases: an evolutionary approach for facing the interpretability problem. *Information* **10**(1), 30 (2019)
23. Senatore, R., Della Cioppa, A., Marcelli, A.: Automatic diagnosis of Parkinson disease through handwriting analysis: a Cartesian genetic programming approach. In: 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), pp. 312–317, June 2019
24. 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)
25. Valenzuela, O., Jiang, X., Carrillo, A., Rojas, I.: Multi-objective genetic algorithms to find most relevant volumes of the brain related to Alzheimer's disease and mild cognitive impairment. *Int. J. Neural Syst.* **28**(09) (2018)