

**Characterization of neurological disorders  
using evolutionary algorithms**

**Chiara Picardi**  
Doctor of Philosophy  
University of York  
Electronic Engineering  
May 2018

## Abstract

The life expectancy increasing, in the last few decades, leads to a large diffusion of neurodegenerative age-related diseases such as Parkinson's disease. Neurodegenerative diseases are part of the huge category of neurological disorders, which comprises all the disorders affecting the central nervous system. These conditions have a terrible impact on life quality of both patients and their families, but also on the costs associated to the society for their diagnosis and management. In order to reduce their impact on individuals and society, new better strategies for the diagnosis and monitoring of neurological disorders need to be considered.

The main aim of this study is investigating the use of artificial intelligence techniques as a tool to help the doctors in the diagnosis and the monitoring of two specific neurological disorders (Parkinson's disease and dystonia), for which no objective clinical assessments exist. The evolutionary algorithms are chosen as the artificial intelligence technique to evolve the best classifiers. The classifiers evolved by the chosen technique are then compared with those evolved by two popular well-known techniques: artificial neural network and support vector machine. All the evolved classifiers are not only able to distinguish among patients and healthy subjects but also among different subgroups of patients. For Parkinson's disease: two different cognitive impairment subgroups of patients are considered, with the aim of an early diagnosis and a better monitoring. For dystonia: two kinds of dystonia patients are considered (organic and functional) to have a better insight in the division of the two groups.

The results obtained for Parkinson's disease are encouraging and evidenced some differences among the cognitive impairment subgroups. Dystonia results are not satisfactory at this stage, but the study presents some limitations that could be overcome in future work.

This thesis is dedicated to my family with love.

# LIST OF CONTENTS

<b>ABSTRACT .....</b>	<b>2</b>
<b>LIST OF CONTENTS .....</b>	<b>4</b>
<b>LIST OF TABLES.....</b>	<b>6</b>
<b>LIST OF FIGURES.....</b>	<b>8</b>
<b>AKNOWLEDGEMENTS.....</b>	<b>10</b>
<b>DECLARATION .....</b>	<b>12</b>
<b>CHAPTER 1: INTRODUCTION .....</b>	<b>13</b>
<b>1.1 NEUROLOGICAL DISORDERS AND PROBLEMS ASSOCIATED.....</b>	<b>13</b>
<b>1.2 ANALYSYS OF THE COSTS ASSOCIATED TO NEUROLOGICAL DISORDERS.....</b>	<b>15</b>
1.2.1 Costs associated to brain disorders.....	15
1.2.2 Costs associated to Parkinson’s disease and dementia.....	19
<b>1.3 HEALTH BURDEN ASSOCIATED TO NEUROLOGICAL DISORDERS .....</b>	<b>22</b>
<b>1.4 MOTIVATIONS OF THE WORK .....</b>	<b>24</b>
<b>1.5 AIMS AND NOVELTY OF THE WORK .....</b>	<b>26</b>
<b>1.6 RESEARCH QUESTION AND WORK ORGANIZATION .....</b>	<b>28</b>
<b>CHAPTER 2: EVOLUTIONARY ALGORITHMS.....</b>	<b>30</b>
<b>2.1 HISTORY.....</b>	<b>30</b>
<b>2.2 EVOLUTIONARY COMPUTATION AND EVOLUTIONARY ALGORITHMS .....</b>	<b>31</b>
<b>2.3 GENETIC PROGRAMMING.....</b>	<b>33</b>
2.3.1 Tree-based representation .....	33
2.3.2 Other representations.....	34
2.3.3 Advantages and problems .....	36
<b>2.4 CARTESIAN GENETIC PROGRAMMING .....</b>	<b>37</b>
<b>2.5 PREVIOUS MEDICAL APPLICATIONS.....</b>	<b>41</b>
<b>2.6 SUMMARY AND CONCLUSIONS .....</b>	<b>45</b>
<b>CHAPTER 3: PARKINSON’S DISEASE.....</b>	<b>47</b>
<b>3.1 HISTORY OF THE DISEASE .....</b>	<b>47</b>
<b>3.2 PARKINSON’S DISEASE SYMPTOMS .....</b>	<b>49</b>
3.2.1 Bradykinesia .....	49
3.2.2 Rest tremor .....	50
3.2.3 Rigidity .....	51
3.2.4 Loss of postural reflexes.....	51
<b>3.3 COGNITION IN PARKINSON’S DISEASE.....</b>	<b>52</b>
<b>3.4 SUMMARY AND CONCLUSIONS .....</b>	<b>54</b>
<b>CHAPTER 4: DYSTONIA .....</b>	<b>55</b>
<b>4.1 HISTORY OF THE DISEASE.....</b>	<b>55</b>
<b>4.2 KINDS OF DYSTONIA SIGNS AND SYMPTOMS.....</b>	<b>56</b>
<b>4.3 CAUSES AND DIAGNOSIS OF DYSTONIA .....</b>	<b>58</b>
4.3.1 Primary dystonia.....	58
4.3.2 Secondary dystonia.....	58
<b>4.4 FUNCTIONAL AND ORGANIC DYSTONIA .....</b>	<b>59</b>
<b>4.5 SUMMARY AND CONCLUSIONS .....</b>	<b>61</b>

<b>CHAPTER 5: REACH AND GRASP EXPERIMENT .....</b>	<b>62</b>
<b>5.1 REACH AND GRASP MOVEMENT .....</b>	<b>62</b>
<b>5.2 METHODOLOGY .....</b>	<b>65</b>
5.2.1 Subjects recruited.....	65
5.2.2 Equipment used .....	66
5.2.3 Experiment description .....	68
5.2.4 Reaching data .....	70
5.2.5 Features extracted .....	73
5.2.6 Grasping data.....	77
5.2.7 Description of the data .....	78
<b>5.3 RESULTS.....</b>	<b>79</b>
5.3.1 Cases analysis.....	79
5.3.2 CGP classification results.....	83
5.3.3 CGP classifier Mathematical expressions and Network Diagrams.....	85
<b>5.4 CONCLUSIONS AND FUTURE WORKS .....</b>	<b>87</b>
<b>CHAPTER 6: WAVELET TRANSFORM ANALYSIS .....</b>	<b>90</b>
<b>6.1 THE CONTINUOUS WAVELET TRANSFORM .....</b>	<b>93</b>
6.1.1 Wavelet energy .....	97
6.1.2 Heisenberg box wavelet transform.....	100
6.1.3 Morlet and Analytic Wavelets .....	103
<b>6.2 THE DISCRETE WAVELET TRANSFORM .....</b>	<b>109</b>
<b>6.3 SUMMARY AND CONCLUSIONS .....</b>	<b>110</b>
<b>CHAPTER 7: HAND OPENING-CLOSING EXPERIMENT .....</b>	<b>112</b>
<b>7.1 AIM AND MOTIVATIONS.....</b>	<b>112</b>
<b>7.2 EMG SIGNAL.....</b>	<b>114</b>
<b>7.3 METHODOLOGY.....</b>	<b>117</b>
7.3.1 Subjects recruited.....	117
7.3.2 Equipment used .....	118
7.3.3 Experiment description .....	119
7.3.4 Hand opening-closing data .....	121
<b>7.4 RESULTS.....</b>	<b>124</b>
7.4.1 Maximum for each region.....	127
7.4.2 Mean for each region.....	129
7.4.3 Standard deviation for each region.....	130
7.4.4 Max, mean and standard deviation for each region.....	132
7.4.5 Fourier Transform comparison.....	134
7.4.6 CGP classifier Mathematical expressions and Network Diagrams.....	135
<b>7.5 CONCLUSIONS AND FUTURE WORK.....</b>	<b>136</b>
<b>CHAPTER 8: CONCLUSIONS AND FUTURE WORK.....</b>	<b>140</b>
<b>8.1 RESEARCH SUMMARY.....</b>	<b>140</b>
<b>8.2 CONCLUSIONS AND FUTURE WORK.....</b>	<b>141</b>
8.2.1 Parkinson’s disease classification.....	142
8.2.2 Dystonia classification .....	143
8.2.3 Final conclusions.....	145
<b>REFERENCES.....</b>	<b>147</b>

## LIST OF TABLES

Table 1. Summary of the class considered with demographic details (Standard deviation, range).....	66
Table 2. Number of samples considered for each class in each condition.....	79
Table 3. SVM and ANN classification results for all conditions. ....	81
Table 4. SVM and ANN classification results for NAT condition. ....	81
Table 5. SVM and ANN classification results for VIS condition. ....	81
Table 6. SVM and ANN classification results for MAX condition.....	82
Table 7. SVM and ANN classification results for MEM condition.....	82
Table 8. Classification results for MEM condition. ....	84
Table 9. Number of subjects and samples considered for each class (organic, functional, controls) considering the muscles FDS and EDC.....	125
Table 10. Classification results considering the case Org&Fun vs Con (inputs MAX).....	127
Table 11. Classification results considering the case Org vs Con (inputs MAX)..	127
Table 12. Classification results considering the case Fun vs Con (inputs MAX)..	128
Table 13. Classification results considering the case Org vs Fun (inputs MAX)..	128
Table 14. Classification results considering the case Org&Fun vs Con (inputs MEAN).....	129
Table 15. Classification results considering the case Org vs Con (inputs MEAN). ....	129
Table 16. Classification results considering the case Fun vs Con (inputs MEAN). ....	129
Table 17. Classification results considering the case Org vs Fun (inputs MEAN). ....	130
Table 18. Classification results considering the case Org&Fun vs Con (inputs STD). ....	131
Table 19. Classification results considering the case Org vs Con (inputs STD). ....	131
Table 20. Classification results considering the case Fun vs Con (inputs STD). ....	131
Table 21. Classification results considering the case Org vs Fun (inputs STD). ....	131
Table 22. Classification results considering the case Org&Fun vs Con (inputs MAX, MEAN, STD). ....	132
Table 23. Classification results considering the case Org vs Con (inputs MAX, MEAN, STD). ....	133
Table 24. Classification results considering the case Fun vs Con (inputs MAX, MEAN, STD). ....	133
Table 25. Classification results considering the case Org vs Fun (inputs MAX, MEAN, STD). ....	133
Table 26. Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for FDS are reported. ....	134

Table 27. Classification results considering the case Org vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for FDS are reported..... 134

Table 28. Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for EDC are reported..... 135

Table 29. Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for EDC are reported..... 135

Table 30. Deep learning test set results. .... 139

## LIST OF FIGURES

Figure 1. Costs per person in Europe in 2010 by type of cost and condition.	16
Figure 2. Total cost by disorder and type of cost in Europe in 2010.	17
Figure 3. Distribution of costs.	18
Figure 4. Total UK cost of individual brain disorders.	18
Figure 5: Direct costs of PD in the UK according to Hoehn and Yahr scale.	19
Figure 6. Average annual cost per person with dementia, by severity and settings (£, 2012/2013 prices).	21
Figure 7. Estimated breakdown of costs of dementia for the UK, 2013.	22
Figure 8. Percentage of total DALYs for selected diseases and neurological disorders.	23
Figure 9. Neurological disorders as percentage of total DALYs for 2005, 2015 and 2030 (estimated) across World Bank income category.	24
Figure 10. Representation of the Evolutionary algorithm's steps.	32
Figure 11. Program tree representing the S-expression (+a(*bc))	34
Figure 12. Arbitrary directed graph with a maximum depth.	37
Figure 13. An example of chromosome's phenotype and genotype.	38
Figure 14. Representation of the chromosome meaning.	39
Figure 15. Effect of a silent mutation.	40
Figure 16. Effect of a single gene mutation on phenotype.	41
Figure 17. Figure copying task, the data are measured using a conventional digitizing tablet. Source: (Smith et al. 2007)	42
Figure 18. Example of two-part velocity feature.	43
Figure 19. ROC curves for the ensemble and its component classifiers.	44
Figure 20. Trace of separations values.	44
Figure 21. Dopamine levels in normal neurons and in Parkinson's affected neurons.	48
Figure 22. 5DT Data Glove 5 Ultra.	67
Figure 23. EM sensor fixed on the wrist under the Velcro strap of the glove.	68
Figure 24. The EM tracking device Polhemus Patriot.	68
Figure 25. Experiment setup.	69
Figure 26. Positional space coordinates (x,y,z).	70
Figure 27. Rotational coordinates (roll,pitch,yaw).	70
Figure 28: Distance profile relative to a trial made by a healthy subject.	73
Figure 29. Distance (upper), velocity (middle) and acceleration (bottom) profiles relative to the same trial of Figure 28.	74
Figure 30. Grasp data relative to a trial performed by a healthy subject.	78
Figure 31. Network diagram of the best classifier evolved for the case PDD vs Con.	87
Figure 32. Scaling and translation of a wavelet plus wavelet transform for a certain scale	92
Figure 34: Dilatation and translation of a wavelet.	95
Figure 35: Wavelet interrogation of the signal.	97
Figure 36. Top: 7 seconds of human ECG containing a defibrillation shock event.	99
Figure 37. The aorta pressure trace (top) with the ECG (middle) and the corresponding wavelet energy plot (bottom) using the Morlet wavelet.	100
Figure 38: Heisenberg boxes in time-frequency domain.	102
Figure 39. Morlet wavelet.	105

Figure 40. Use of the Morlet wavelet to detect an abrupt change in a signal periodicity.	106
Figure 41. Examples of generalized Morse wavelets.	108
Figure 42. Parameter space for the generalized Morse wavelet superfamily.	109
Figure 43. Model of the single muscle fiber detection by using a differential measure of muscle activity.	115
Figure 44. Model of the triphasic motor unit action potential.	116
Figure 45. Electromagnetic sensor attached to thumb and index fingers.	118
Figure 46. Shimmer EMG sensors fixed on the skin in correspondence of the selected muscles.	119
Figure 47. Shimmer EMG unit secured to the arms with elasticated straps during the movements.	119
Figure 48. Location of the muscles on the arm.	121
Figure 49. Thumb-index fingers distance on y-axis, time on x-axis.	122
Figure 50. EMG data of a patient with the corresponding CWT scalogram.	123
Figure 51. Grid built on the spectrogram.	124
Figure 52. Network diagram of the best classifier evolved for the case Org vs Con considering as inputs the 14 maxima extracted by the FDS scalogram grid.	136

## AKNOWLEDGEMENTS

I am extremely grateful to both my supervisors - Prof Steve Smith and Dr David Halliday – for giving me the opportunity to work on this interesting project, guiding me with their wise suggestions.

I thank Steve for believing in me from the beginning, for all support that he gave me along the way, but overall for the huge contribution to my professional growth. I will be eternally grateful to Steve for convincing me to accept the PhD offer. I choose that I will never regret and that permitted me to know York, city that changed my life, and to undertake this research.

I thank David for his precious advices, his continuous support and his availability. I thank him also for revising part of my thesis providing insightful comments to improve the grammar and the content of this thesis.

I thank Dr Jeremy Cosgrove and Dr Rachel Newby for conducting the experiments, recording the data used in this thesis.

I thank Jeremy for his support at the beginning of this journey, for his patience in explaining me all the medical terms and for his availability to clarify all my doubts. We made a wonderful team and I am glad that I had the opportunity to work with him.

I thank Rachel for the measurements and for her availability when I needed clarifications.

I thank Mirella Aquila for revising the grammar of the thesis. Thanks Mirella, you did a great and invaluable job.

I thank Amir and Siti for all their support along the journey and their precious advices. Thank you guys I certainly will miss you both.

A special thank is for the best office mate in the world: Fady Medhat. I thank Fady for all the laughs, for the support during hard times and for opening my eyes many times during our research discussions. Thanks Fady, the office was never the same after you left.

I thank all my York friends and in particular - Rosaria, Manal, Rachele and Yasmin – for all their support. I thank Rosaria for being ever here for me, for all her support during the ups and downs and for all the meals and beers at the pub. I thank Manal for being ever on my side, for teaching me that the differences do not exist and for her kindness. I thank Rachele for our endless chats, for the funny evenings and for

her precious advices. I thank Yasmin for her support during the hard times and for being here when I need.

I thank Rosie my helper and my friend for making my life easy. I thank her also for tolerating me during the hard times.

I thank my friends Marco and Maria for all the times that they supported and tolerated me during the times of crisis that have beset this study. Thanks to you both for believing in me, for your encouragements, support and for being on my side every time.

I thank Marina for our endless telephone chats, supporting each other during our respectively researches. Thanks to you Marina for being ever available during my hard times, our friendship erases the physical distance.

I thank my boyfriend Robert for his patience during the hard times, for his support and for his constant encouragement. Thanks Rob for being ever on my side.

The final and biggest thank is for my family: my mother Pina, my father Ciro and my twin sister Marta. Without the encouragement and the support of my family I would never have been able to complete this difficult journey. I thank my family for supporting me every time, for believing in me also in the hard times and for erasing the physical distance being available for me every time I need. This work is entirely dedicated to my family hoping to make them proud of me.

## **DECLARATION**

Some of the research presented in this thesis has previously been published by the author (Picardi et al. 2017). I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

# CHAPTER 1: INTRODUCTION

Nowadays, thanks to scientific progress and the discovery of new drugs, people live longer and, as a consequence, there are more elderly people than a few decades ago. The process of becoming older leads to new challenges for society. In fact there are a lot of neurodegenerative age-related diseases likely to become more common, due to increased life expectancy. These conditions not only have a terrible impact on patients and their family life, but also on the costs of the society for diagnosing, monitoring and managing these conditions. Neurodegenerative diseases are part of the huge category of neurological disorders which comprises all the brain disorders affecting the central nervous system. In this work we will investigate the application of artificial intelligence techniques as a tool to help in the diagnosis and monitoring of two specific neurological disorders, for which no objective clinical assessments exist. The aims of the techniques presented are: reducing the costs for diagnosing and monitoring specific diseases with a simple and cost efficient test; improving the patients and families life with an early and objective diagnosis and also designing a better target therapy identified by the improved monitoring and diagnosis.

This chapter describes the motivations, the aims and the novelty of this study. It is divided in six main sections. In section 1.1 a brief introduction about the incidence and the costs of neurological disorders is given. Section 1.2 describes the costs related to neurological disorders, examining graphs and tables. Section 1.3 discusses health burden associated with neurological disorders by estimating the impact on quality life. The motivation of the research along with an overview of its aims is given in section 1.4. In section 1.5 the aims of the work are described in details along with the motivations of the work novelty. The study is summarized as a research question in section 1.6, where also the thesis organization is given.

## 1.1 NEUROLOGICAL DISORDERS AND PROBLEMS ASSOCIATED

In the last decades medical advances and improvements in sanitation led to an increased life expectancy. The world-wide life average reached about 70 years in 2014 (six years longer than in 1990) and about 80 years considering only the developed countries (in the early 20<sup>th</sup> century this was about 50 years) (Jin et al. 2015). In a recent study Kontis and colleagues predicted that life expectancy could increase more in the next years, reaching in 2030 an average of 85 years or more in the industrialised countries (Kontis et al. 2017). In future years then, a further increasing of the life average is expected, leading to more people living longer.

Although the increased life average is a positive aspect, it leads to new challenges. Getting older, in fact, is strictly linked to biological and cognitive degeneration resulting in a consequent increasing of neurodegenerative diseases (Jin et al. 2015). Neurodegenerative diseases such as Alzheimer's and Parkinson's disease are

known to be strongly age related, there is no cure, they are difficult to be slowed significantly and usually end with death. Parkinson's disease is the second most common age related disease after Alzheimer's disease, having a prevalence of approximately 0.5 to 1 percent for those aged between 65-69, and rising to 1 to 3 percent for those aged 80 and above (Nussbaum and Ellis 2003). Ageing decline appears to be important for the development of Parkinson's disease itself, leading to the loss of neurons related to the disease (Reeve, Simcox and Turnbull 2014).

Neurodegenerative diseases are part of the more general category of brain disorders. Brain disorders can be very disabling and often with the progression of the disease the person is unable to perform daily activities. There are a lot of different brain disorders: some of them affect only the psychological functions without directly affecting the central nervous system (mental disorders) such as sleep disorders, anxiety disorders etc. , while others affect the central nervous system (neurological disorders), affecting also muscles (i.e. neuromuscular disorders) or movements (i.e. neurodegenerative diseases) and resulting often in very disabling conditions. In this work we are interested in neurological disorders. These incurable and disabling conditions have a devastating impact on the individual, family and society. Many people affected by neurological disorders are not able to continue their life independently relying on the family, who are required to find a way to assist them.

Neurological disorders then have a terrible impact on the quality of life of the people affected and their family, but also have a financial burden for the family and society. In order to compute the costs associated to these conditions, we need to take in account not only the cost of treatment itself, but also the loss of productivity of patients and their caregivers which are exposed to an emotional and financial burden. Many caregivers have to leave their job and all leisure activities to look after their family member affected by the disease or have to pay someone to care for him/her while they are at work. The cost of brain disorders (comprising neurological disorders) in Europe was estimated as €798 billions in 2010 of which 60% was related to the direct costs while the 40% to lost productivity (Gustavsson et al. 2011). The cost of Parkinson's disease for the society was estimated in eleven different countries (Austria, Czech Republic, Finland, France, Germany, Italy, Portugal, Russia, Sweden, UK and US) between 1998 and 2011 (Mateus and Coloma 2013). The results took in account both the direct and indirect costs. The direct costs included the cost of the illness itself (treatment, medical analysis, hospitalization, etc.) while the indirect costs comprised all the cost derived by the illness (loss of productivity of patient and caregiver, possible early retirement etc.). The authors concluded that, although different countries had different results, the costs associated to the disease increases with the disease progression which cause more need for healthcare or caregiver support affecting more the patient's ability to perform everyday tasks. The costs computed for each country, with different methodologies, evidenced the health and economic burden for the society (e.g. annual cost for a PD patient in UK was estimated as £1.4 billion). The estimation of the indirect costs was the most difficult part of the study, but these costs are

important because they represent a significant burden for patients, caregivers and society as whole.

In another study Wimo and colleagues estimate the worldwide economic impact of dementia in 2010 (Wimo et al. 2013). Dementia is a very disabling neurodegenerative disease that affects everyday tasks, making the subject unable to live independently. The results of Wilmo's study evidenced a worldwide cost of US\$604 billion in 2010. 70% of the costs was found in western Europe and North America where social care and informal care (care from a family member) contributed in almost equal manner to the overall cost. In low and middle-income countries the contribution of informal care was higher respect to social care one which contribution was negligible.

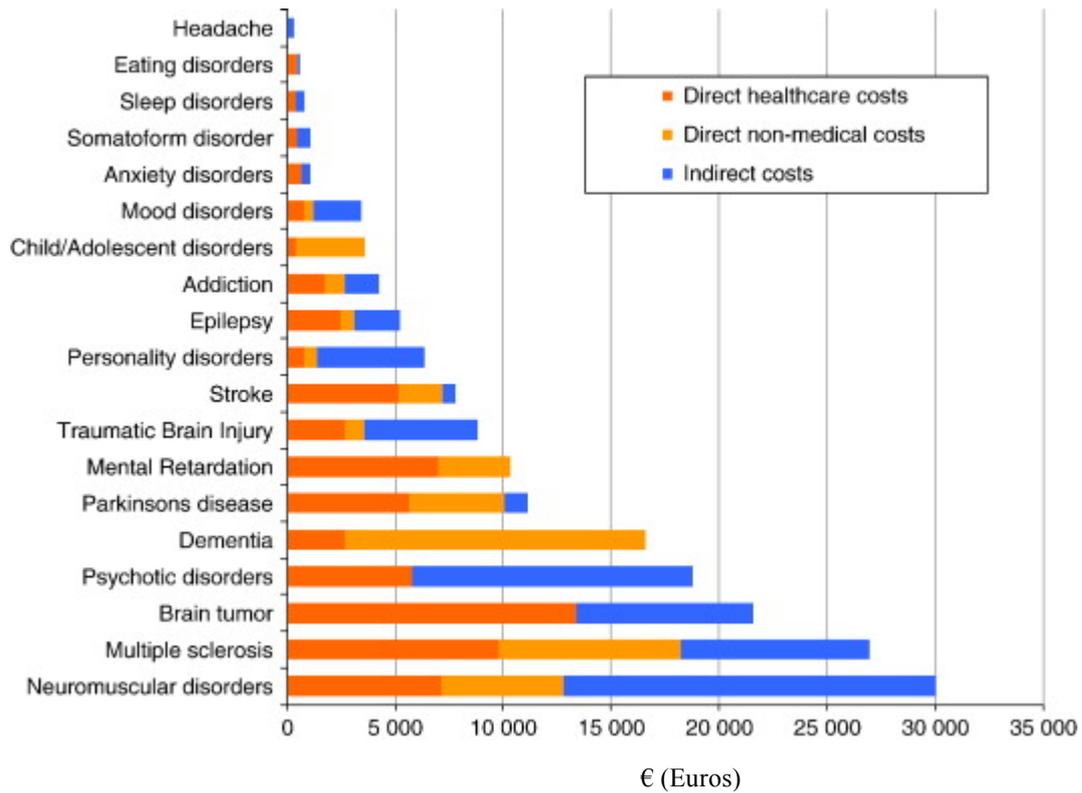
The overview of this section gives an idea of the diffusion and the burden associated to neurological disorders. The problem for society will increment with the increasing of life expectation, predicted to increase every year (Kontis et al. 2017). In the next sections we will analyse in details the two major problems associated to neurological disorders: the costs and the aggravation of the quality of life.

## **1.2 ANALYSYS OF THE COSTS ASSOCIATED TO NEUROLOGICAL DISORDERS**

In the previous section we examined some findings showing the increasing of incidence of neurodegenerative diseases, due to the increasing of life expectancy. The costs associated to dealing with more people affected by these extremely disabling conditions are high for society and they are destined to increase with the predicted life expectancy. In this section the costs associated to these conditions and more in general to the neurological disorders are examined in details. The section is divided in two subsections: firstly the costs associated to brain disorders are examined; secondly we examine the costs associated to two particular neurodegenerative diseases: Parkinson's disease and dementia.

### **1.2.1 Costs associated to brain disorders**

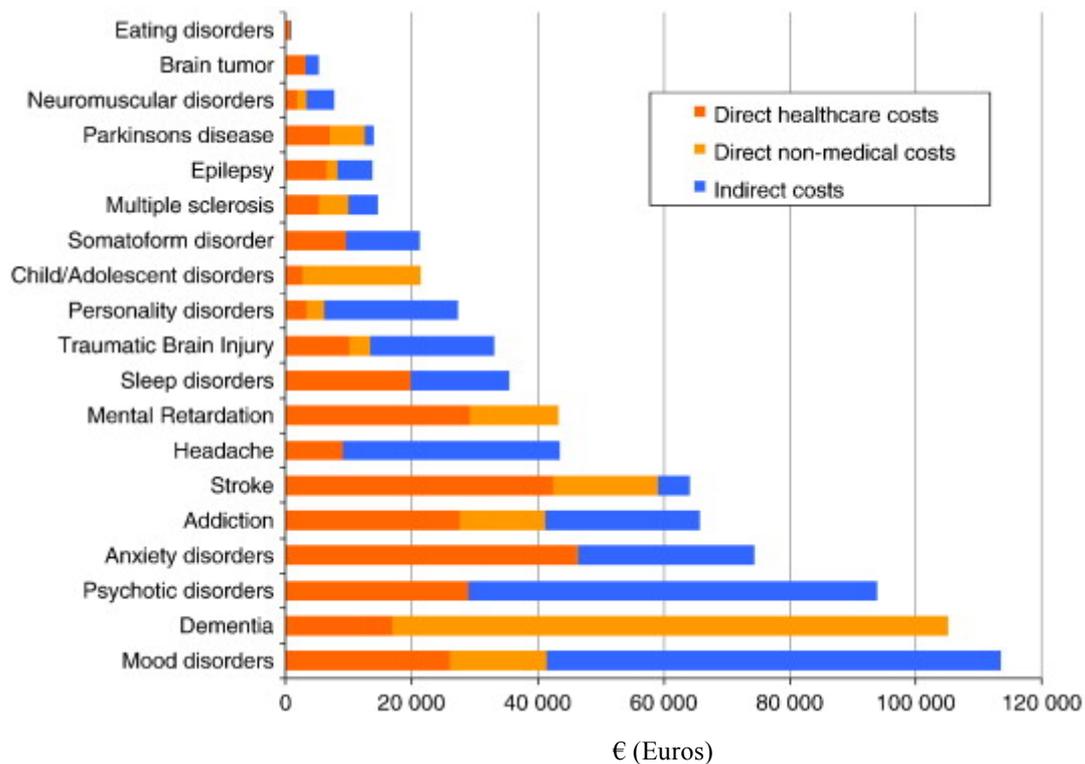
A study by Gustavsson and colleagues estimated costs for brain disorders in 2010 (Gustavsson et al. 2011). The costs reported are divided in: direct healthcare costs (relative to treatment and diagnosis of the disease), direct non-medical costs (relative to goods and services associated to the disease e.g. social services, special accommodation, and informal care), and indirect costs (e.g. loss of productivity due to work absence or early retirement). In figure 1 the annual costs per person in Europe in 2010 are summarized.



**Figure 1. Costs per person in Europe in 2010 by type of cost and condition.**  
 The different conditions are reported on y-axis while the costs associated are reported on x-axis. Source: (Gustavsson et al. 2011)

Neuromuscular disorders, such as dystonia considered in this work, had the highest costs equal to 30000€ per person in 2010, with a prevalence of indirect costs. Dementia had also high costs around 17000€, with prevalence of direct non-medical costs and without presence of indirect costs (due probably to the age of the subjects affected). Parkinson’s disease costs instead were estimated around 12000€ per person, with a prevalence of direct medical costs due to the diagnosis and the treatment of the disease.

In figure 2 the total costs per disease spent in Europe in 2010 are reported. Figure 2 is simply derived from figure 1 aggregating the costs and considering the people affected by the different diseases. The total costs are the highest for mood disorders followed by dementia. Surprisingly neuromuscular disorders are not in the most elevated costs as in figure 1. This is due to the fact that the people affected by neuromuscular disorders are considerably less than those affected by the other diseases.



**Figure 2. Total cost by disorder and type of cost in Europe in 2010.**  
**The different conditions are reported on y-axis while the costs associated are reported on x-axis.**  
**Source: (Gustavsson et al. 2011).**

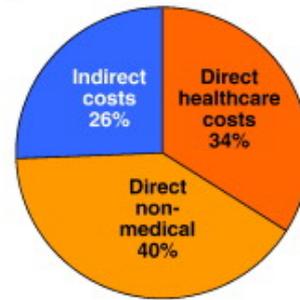
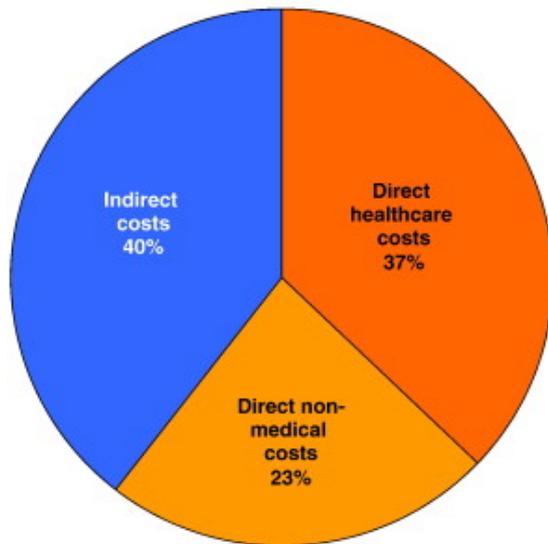
In figure 3 the percentage of the various costs in brain, mental and neurological disorders are shown. For brain disorders in general the direct costs were estimated around 60% while the indirect costs only about 40%. For mental disorder the highest costs were the indirect ones estimated at 48%, while for neurological disorders the highest percentage were for direct costs (74%). These results are predictable because mental disorders cause a major loss of productivity than neurological disorders affecting usually elderly people. We need to consider also that informal care is included in the direct non-medical costs and this component is very significant in neurological disorders. In fact often people affected by neurological disorders are unable to perform daily tasks relying on a family member as an informal carer.

This estimation of the different costs in 2010 in Europe gave an idea of the elevated costs associated to brain diseases in general. Interestingly, despite the neuromuscular disorders had the larger costs per person that year (figure 1); they do not seem having the same large impact on the total costs (figure 2).

The same costs were estimated in 2010 considering only UK (Fineberg et al. 2013) and the results are reported in figure 4.

A) Total disorders of the brain

B) Neurological disorders



C) Mental disorders

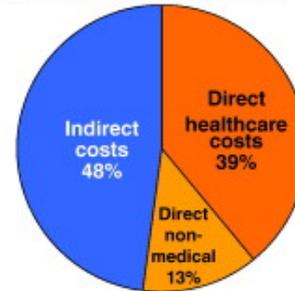
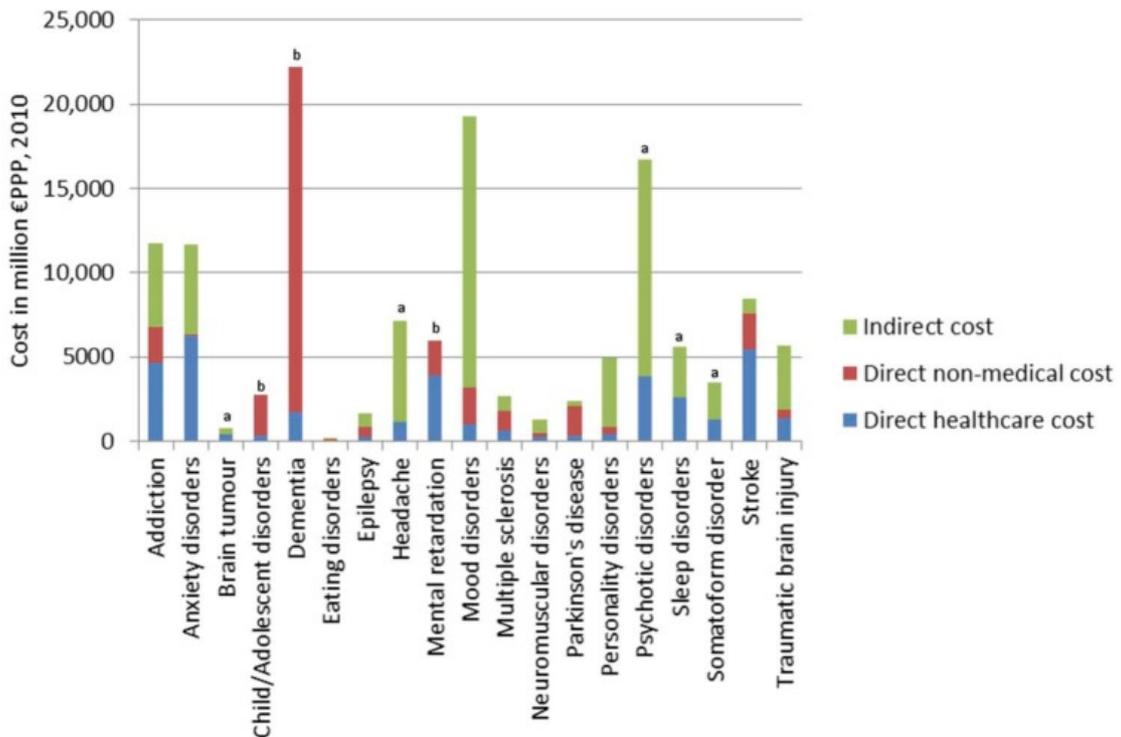


Figure 3. Distribution of costs.  
Source: (Gustavsson et al. 2011).

### Total Cost - United Kingdom



<sup>a</sup>Missing data for direct non-medical costs;

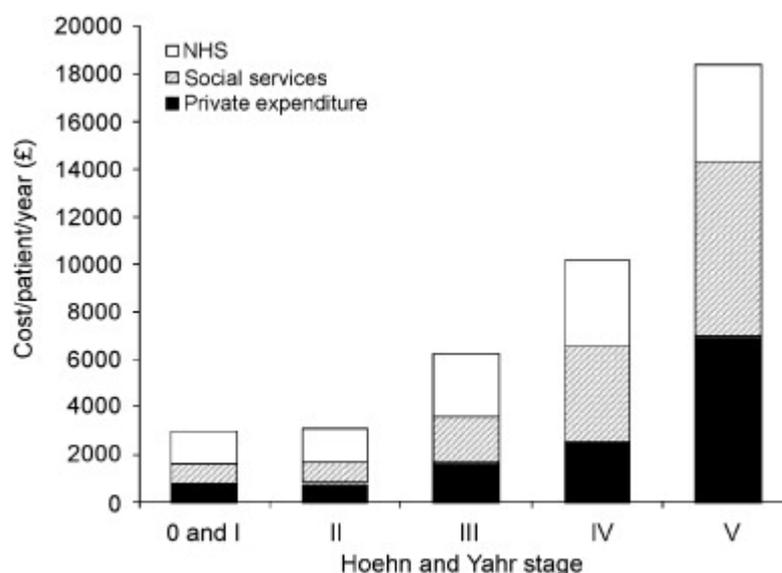
<sup>b</sup>Missing data for indirect costs

Figure 4. Total UK cost of individual brain disorders.  
On x-axis the different condition are reported while on y-axis the costs in euros are reported.  
Source: (Fineberg et al. 2013)

In UK the highest costs are found for dementia followed by mood disorders. Also in this case the costs associated to the neuromuscular disorders are extremely low considering that the cost per person is the highest.

### 1.2.2 Costs associated to Parkinson’s disease and dementia

Parkinson’s disease (PD) is an age-related disease, very common and intended to become ever more common with the increasing of life expectancy. In 2007, Findley estimated the total cost in UK being between £449million and £3.3billion annually (Findley 2007). The costs are divided in direct costs (healthcare resource use and drugs) and indirect costs (mortality costs, lost productivity and carer replacement costs). The direct costs are divided also in: NHS costs, social service costs and private PD-related expenditures (private care, equipment, travel etc.). In figure 5 the direct costs of PD in the UK according to Hoehn and Yahr scale (Hoehn Mm Fau - Yahr and Yahr 1967) are reported. Hoehn and Yahr scale is a quantification system to evaluate how the PD symptoms progress. In the scale there are five stages measuring the different grade of disability in a crescent way. Patients in the last stage usually are wheelchair users or bedridden unless aided.



**Figure 5: Direct costs of PD in the UK according to Hoehn and Yahr scale.**  
**Source: (Findley 2007).**

In figure 5 we can acknowledge the increasing costs associated to the progression of the disease, the findings confirmed later by another worldwide study (Mateus and Coloma 2013). The progression of the disease leads to an increasing disability resulting in higher costs for caring. During the progression stages, in fact, the patients could develop a cognitive impairment which could be mild or severe. In this work two cognitive impairments are considered: mild cognitive impairment and dementia. Mild cognitive impairment is less severe than dementia, it affects the memory but it is not too severe to affect the ability to live independently. Dementia

is a severe cognitive impairment leading to inability to perform daily tasks and then resulting in a very disabling condition.

In a report, published in 2014, Prince and colleague reported the costs associated to dementia in UK relative to 2013 (Prince et al. 2014). This report is an update of a previous report published in 2007 (Knapp et al. 2007) . In figure 6 the annual costs for each person affected by dementia in UK is reported. The results are divided according three different severity levels of dementia (mild, moderate and severe) and two different settings (community and residential care). For people living in community the following annual healthcare costs are found: £2,751 for those with mild dementia, £2,695 for those with moderate dementia and £11,258 for those with severe dementia. People with severe dementia, in the community, have considerable higher costs respect to the two other categories.

In the residential care setting the situation is different, with the following costs: £4,504 (mild), £9,438 (moderate) and £8,689 (severe). This change is due to the fact that NHS pays full costs or in some case part of the costs for the people staying in nursing home care. The healthcare annual average costs per person is estimated to be £5,285 which approximately £650 are spent for the diagnosis.

The average annual social care costs per person, reported in figure 6, for people living in the community are: £3,121 (mild dementia), £7,772 (moderate) and £10,321 (severe). The increasing of the costs with the severity of the disease confirms the same trend of the healthcare costs. For people residing in nursing homes the costs are higher, but there is not a big difference among the different severity levels: £24,737 (mild), £25,715 (moderate) and £25,874 (severe). Considering all the severity levels and both settings the average annual social care costs of people with dementia is £12,584.

	Healthcare	Social care	Unpaid care	Other costs	Total costs
<b>People with dementia living in the community (average cost)</b>					
Mild dementia	2,751	3,121	19,714	137	25,723
Moderate dementia	2,695	7,772	32,237	137	42,841
Severe dementia	11,258	10,321	33,482	136	55,197
All severity levels	3,152	4,054	21,956	137	29,298
(Sector cost as % of total)	(10.8%)	(13.8%)	(74.9%)	(0.5%)	(100%)
<b>People with dementia living in residential care (average cost)</b>					
Mild dementia	4,504	24,737	1,067	136	30,444
Moderate dementia	9,438	25,715	2,901	136	38,190
Severe dementia	8,689	25,874	2,119	136	36,817
All severity levels	8,542	25,610	2,450	136	36,738
(Sector cost as % of total)	(23.3%)	(69.7%)	(6.7%)	(0.4%)	(100%)
<b>All settings (average cost)</b>					
Mild dementia	2,932	5,362	17,781	137	26,212
Moderate dementia	7,837	21,455	9,865	136	39,294
Severe dementia	9,300	22,176	9,575	136	41,187
All severity levels	5,285	12,584	14,237	136	32,242
(Sector cost as % of total)	(16.4%)	(39.0%)	(44.2%)	(0.4%)	(100%)

**Figure 6. Average annual cost per person with dementia, by severity and settings (£, 2012/2013 prices).**  
**Source: (Prince et al. 2014)**

The average annual unpaid care costs per person for people in the community are: £19,714 (mild), £32,237 (moderate), and £33,482 (severe). The trend of increasing costs with severity of the disease is confirmed also in this case. While considering the residential care setting obviously the unpaid care costs are considerably lower than the community setting: £1,067 (mild), £2,901 (moderate) and £2,119 (severe). The overall average annual unpaid care costs considering all the severity levels and both settings is £14,237 per person.

The column headed “Other costs” comprises: costs for missing persons, advocacy and support costs, research costs and costs of premature mortality (loss of productivity, earnings, etc.). These costs are negligible with respect to the three previous examined costs (health care, social care and unpaid care).

In patients living in the community we notice that unpaid care accounts for three-quarter of the total costs (74.9%). Unpaid care in the community settings has a major contribution on the total costs, followed by: social care (10.8%), healthcare (13.8%) and other costs (0.5%).

In a residential care setting the situation is different. Social care has a major contribution on the total cost (69.7%) followed by healthcare (23.3%), unpaid care (6.7%) and other costs (0.4%).

In figure 7 an estimation of the costs of dementia for the UK in 2013 is reported. The estimation is done considering the costs for person reported in figure 6 and

projecting these costs on all people affected by dementia in 2013. The overall cost computed aggregating all the single costs is estimated to be £26.3 billion, of which the unpaid care element accounts for 44.2%, the social care element accounts for 39%, the healthcare element accounts for 16.4% and the other costs element accounts for 0.4% (figure 5, all settings). These findings evidence the annual large cost for dementia in UK intended to increase as the life expectancy increases.

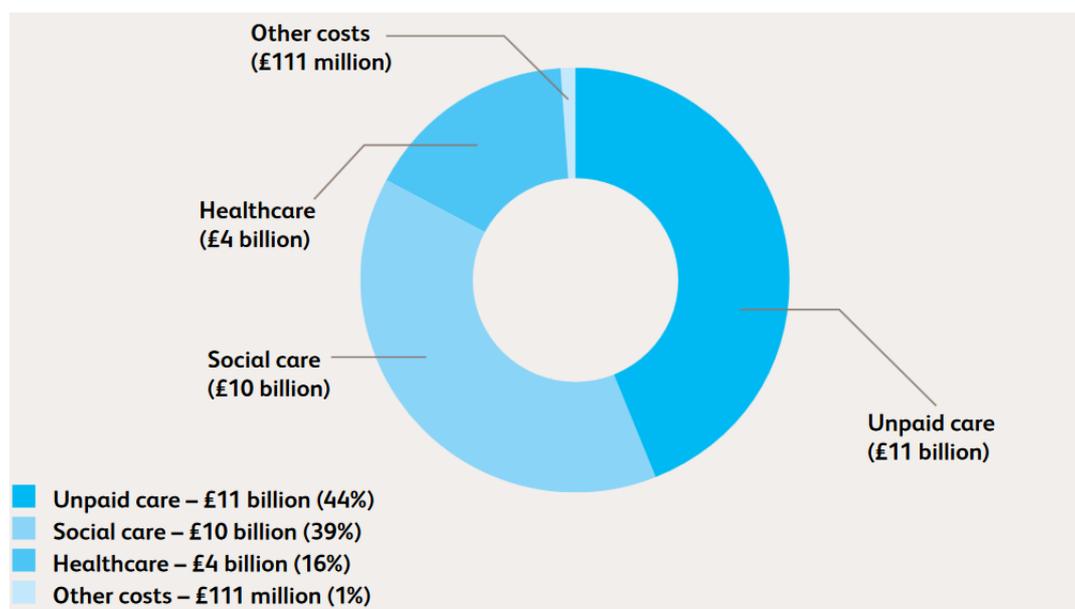


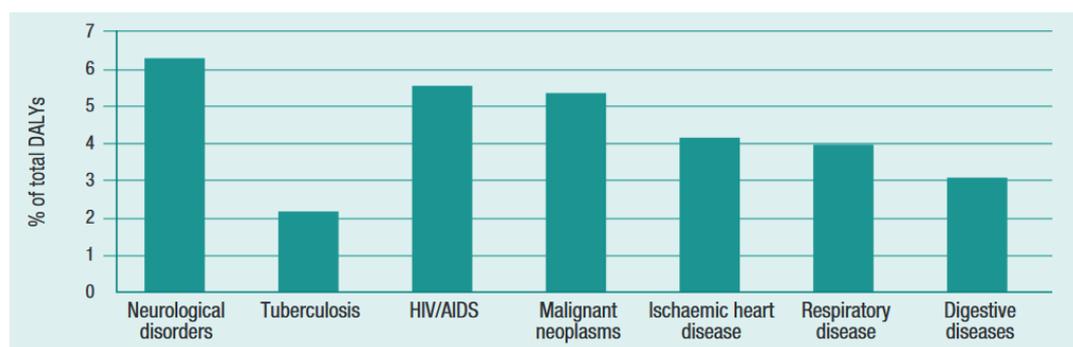
Figure 7. Estimated breakdown of costs of dementia for the UK, 2013.  
Source: (Prince et al. 2014)

### 1.3 HEALTH BURDEN ASSOCIATED TO NEUROLOGICAL DISORDERS

Neurological disorders cause premature deaths and health issues resulting often in disability which affects badly the quality life of the patients and their family. In order to assess the health burden associated to these terrible conditions we need to define a measure able to estimates not only the premature deaths, but also the impact of the disability on the quality life. A time-based metric was defined as the sum of the years of life lost because of premature mortality (YLL) and the years of healthy life lost as results of disability, weighted by the severity of the disability (YLD) (World Health 2006). This metric was called disability-adjusted life years (DALYs) and was used in 1993 to assess the global burden of the disease (GBD) for the year 1990 (Murray, Lopez and Jamison 1994, Murray, Lopez and World Health 1996, Lopez and Murray 1998). DALYs estimates the number of years in which the patient modify his life for the disability associated to the condition equivalent to the number of healthy life years lost. One DALY is then equivalent to one lost year of healthy life.

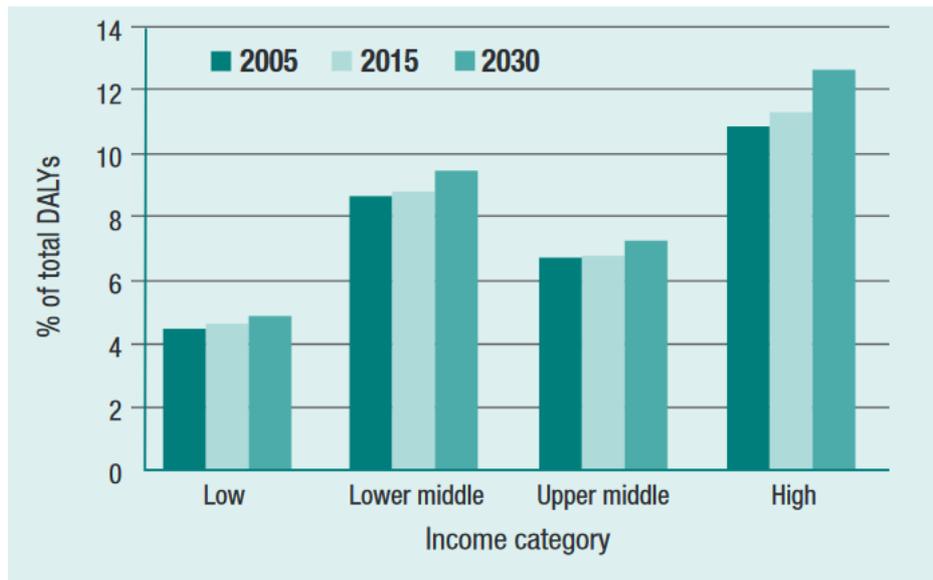
In 2006 the World Health organisation published a report which estimated the global burden of neurological disorders using the DALYs measure (World Health 2006). Some of the findings of that report are reported in this section to estimate the impact of neurological disorders on the quality life.

The percentage of total DALYs for neurological disorders is compared to those for selected diseases in figure 8. The selected diseases are neuropsychiatric disorders plus diseases belonging to other categories. Neurological disorders have the greatest contribution on a global burden of disease reaching over 6% of total burden.



**Figure 8. Percentage of total DALYs for selected diseases and neurological disorders.**  
**Source:(World Health 2006)**

The percentage of total DALYs for neurological disorders in 2005 and its estimation in 2015 and 2030 is reported in figure 9. The results are divided for income category based on World Bank estimates of gross national income (GNI) per capita in 2001 (The World Bank 2003). Each country is classified as low income (GNI US\$ 745 or less), lower middle income (GNI US\$ 746–2975), upper middle income (GNI US\$ 2976–9205), and high income (GNI \$ 9206 or more). The contribution of neurological disorders on global burden increased each year (figure 9). For high income countries we have the highest contribution of neurological disorders on global burden, this is probably due to the greatest life expectancy.



**Figure 9. Neurological disorders as percentage of total DALYs for 2005, 2015 and 2030 (estimated) across World Bank income category. Source: (World Health 2006)**

The findings discussed gave an idea of the health burden associated to neurological disorders in the world. The DALYs measure comprises both the years lost for a premature death and the healthy life years lost by disability. Such disability has a terrible impact on patients' life and also on their family and friends. In fact when the patient is unable to perform daily tasks, a family member or a friend becomes often a full-time informal carer, and her/him may have to leave his or her job and all leisure activities. The carer has a big emotional burden that sometimes becomes also a financial burden, when there is the necessity to pay someone else to look after the patient. The quality life of patients with neurological disorders and their families need to be improved by reducing the impact of the disability. In the next section we will talk about how this work can help in improving patients' lives and, at the same time, in reducing the associated costs.

## **1.4 MOTIVATIONS OF THE WORK**

In the previous two sections (sections 1.2 and 1.3) we discussed the incidence of neurological disorders, the cost associated and the health burden due to the disability caused by these conditions. Particularly in section 1.2 some graphs evaluate the costs of neurological disorders in UK but also across Europe. High costs were found for neurological disorders with dementia (neurodegenerative disease) having one of the highest cost for both Europe and UK in 2010 (figures 2 and 4). The costs in Parkinson's disease instead are shown to increase with the progression of the disease (figure 5). Another important aspect is the quality of life associated to these terrible conditions. An estimation of the quality of life associated to neurodegenerative disorders along with their contribution on global diseases burden, is given in section 1.3. In order to estimate the health burden of neurological disorders, the DALYs measure is used, comprising both the number of

years lost for premature death and healthy life years lost for the disability. The elevated costs and the impact on the patients quality of life evidence the need to improve the diagnosis and the management of these conditions in order to reduce the costs and improving the life of the patients and their families. This need becomes even more necessary thinking that, the incidence of the neurodegenerative diseases which are age-related such as dementia is intended to increase due to the predicted increasing of life expectancy. Research must focus on finding better ways to diagnose, monitor and treat neurological disorders. The discovery of new drugs can reduce the costs, improving the quality of patients' lives. From a medical point of view, it is important that society invests money in the research of new drugs, improving treatments and hopefully finding a cure to these disabling conditions. Therefore the diagnosis and the monitoring of the neurological disorders are also crucial. With the correct diagnosis we can better target the care and the treatment of the disease minimizing the costs. The diagnosis is a crucial part and efforts have to be made in finding better and less expensive ways to obtain a more objective diagnosis. The better target therapy could be beneficial for the subjects, improving their quality of life. On the other hand monitoring is also very important because it permits clinical practitioners to evaluate the progression of the disease modifying medications and treatment, but also permits the evaluation of new drugs as soon as they become available. The modification of the treatment could improve patients' quality life with a better management of the condition.

In this work we consider two neurological disorders: Parkinson's disease (neurodegenerative disease) and dystonia (neuromuscular disease). The basic idea of the research is to use artificial intelligence to provide a better insight in these conditions with the aim of improving diagnosis and monitoring. Parkinson's disease is an age-related disease where instances will increase with the increasing life expectancy. Dystonia was chosen because there are many difficulties in its diagnosis particularly when it is functional dystonia (section 4.4). The prevalence of dystonia was estimated to be around 100-150 per million in some studies conducted in Europe and in Japan (Matsumoto et al. 2003) . Although the precise incidence and prevalence of dystonia is actually unknown (Steeves et al. 2012). Steeves and colleagues evidenced the underestimation of people affected by dystonia, due to the difficulties to diagnose some kind of dystonia manifesting in different ways. The authors remarked of the need to find specific dystonia features in subjects affected by a movement disorder, in order to reduce the misdiagnosis rate. Dystonia is part of neuromuscular diseases that had the highest costs per person in Europe in 2010 (figure 1), considering the underestimation the total costs shown in figure 2 could be higher than represented.

Parkinson's disease affects approximately one person in every 500, equivalent to about 127,000 people in UK (Website of the parkinsons's disease society 2016). The diagnosis actually is based on subjective clinical assessment resulting in up to 25% percent of patient diagnosed with PD found to have other similar condition (Playfer 1997). Many patients also develop cognitive impairment with different severity grades. This cognitive impairment need to be detected and monitored to

minimise the risk of developing dementia (Pedersen et al. 2013), with a consequent reduction of costs and an improvement in quality of life. New methods for a more objective diagnosis and also to identify early the patients presenting cognitive impairments may improve the costs and be more reliable. The first aim of this research is to find a new objective method to help doctors in the diagnosis of the Parkinson disease and also in detecting and monitoring the possible cognitive impairment.

Two kinds of different dystonia are considered: organic and functional dystonia (section 4.4). The distinction between these two categories is not clear. The diagnosis of organic dystonia is based on the evaluation of its core motor features and temporal evolution (Albanese et al. 2013) with a current misdiagnosis rate of 25-52% (Pal 2011). On the other hand the diagnosis of functional dystonia is based on assessment of inconsistency and incongruence (in both the history and neurological examination) with organic disease patterns (Espay and Lang 2015, Ganos, Edwards and Bhatia 2014), with a large risk of diagnosing atypical forms of organic dystonia as functional ones. The second aim of this research is to find a more objective and cost efficient method for the diagnosis of organic dystonia which is also able to detect the differences between functional and organic dystonia.

In summary Parkinson's disease and dystonia are chosen because they are two neurological disorders for which no objective clinical assessment currently exist and their diagnosis is based on subject clinical assessment that is often no very reliable. The need of a more objective diagnosis for both of them is evidenced by the existing misdiagnosis rate equivalent approximately to 25% for Parkinson's disease (Playfer 1997) and between 25-52% for dystonia (Pal 2011). The classifier evolved with Cartesian Genetic Programming can be used as a tool of support for the doctor in the diagnosis, designing a simple and cost efficient test. During the disease it can be used for monitoring and diagnosing the cognitive impairment associated to the Parkinson's disease. Ideally when the doctor suspects a cognitive impairment can use the test to detect it and also monitor the improvement of the cognitive impairment (e.g. due to the new drugs) with the same test.

## **1.5 AIMS AND NOVELTY OF THE WORK**

This work investigates the use of artificial intelligence techniques, such as evolutionary algorithms (chapter 2), to classify and monitor neurological conditions through assessment of movement disorders and cognitive decline using standard clinical tests. The potential of evolutionary algorithms in improving diagnosis and monitoring of neurological disorders will be investigated. The improvement of diagnosis and monitoring have a beneficial impact on both the costs and the quality of life associated, as mentioned in section 1.4.

Two neurological disorders are selected: Parkinson's disease and dystonia. Previous studies investigated the use of the evolutionary algorithms in the classification of Parkinson's disease patients (Smith et al. 2007, Lones et al. 2014,

Lacy et al. 2013) but none before investigated their use in the detection and monitoring of cognitive decline. The characterisation of dystonia using evolutionary algorithms, to the author's knowledge, has not previously been investigated and currently no objective clinical assessment exists (as in Parkinson's disease).

In this work then there are two main aims:

1. Finding a classifier able to distinguish between Parkinson's disease patients and healthy subjects, but also among different subgroups of patients which present different cognitive impairments.
2. Finding a classifier able to distinguish between dystonia patients and healthy subjects but also among two different subgroups of patients (organic and functional ones).

In order to achieve the first aim, a simple reach grasp experiment was performed. All the reaching and grasping kinematic data were collected respectively with an electromagnetic sensor placed on the wrist and with a special computer data glove (description in chapter 5). These data were used to compute different features used as inputs of the evolutionary algorithms to evolve the best classifiers (chapter 5).

In order to achieve the second aim, we considered an experiment comprising several upper limbs movements. During the experiment the activity of four different muscles (EMG data) located on the arm was recorded. Also the kinematic data were recorded with two electromagnetic sensors: one placed on the thumb and the other on the index finger. Hand opening-closing was selected among the movements and the analysis was performed only considering this movement. The two main reason of selecting hand opening-closing were: the major activity of the muscles recorded respect to other movements and the easy way to divide the single hand opening-closing cycle using the kinematic data available. The EMG data were pre-processed using the wavelet transform (chapter 6) and then different features were extracted and used as inputs of the evolutionary algorithms to evolve the best classifiers (chapter 7).

The study is novel for the following reasons:

1. The cognitive decline in Parkinson's disease is assessed examining the kinematic data collected from a simple reach and grasp experiment.
2. A set of features is extracted from the kinematic data collected in the reach and grasp experiment. This set was computed merging different features obtained in two previous studies (Caselli et al. 1999, Alberts et al. 2000). The set of features is then used as inputs to the evolutionary algorithms to evolve the best classifiers able to distinguish between Parkinson's disease patients and healthy subjects but also among different cognitive subgroups of patients.

3. The application of movement sensing technology and evolutionary algorithms in order to provide novel insights into organic, as well as functional, dystonia.
4. The continuous wavelet transform (chapter 6) is used to pre-process EMG dystonia data before extracting the features given as inputs of the evolutionary algorithms.
5. An artificial intelligence technique is used to classify dystonia patients and healthy subjects but also organic and functional patients.

## **1.6 RESEARCH QUESTION AND WORK ORGANIZATION**

In section 1.4 the aims of the work and the reasons of its uniqueness are summarized. In this section all the work is summarized with a research question or hypothesis. In the conclusions (chapter 8) the research question will be revisited.

The work can be summarised by the following research question:

***“Can Evolutionary algorithms provide a means for monitoring and diagnosing of specific neurological disorders?”***

In this study we will investigate the research question concentrating on two specific neurological disorders: Parkinson’s disease and dystonia. For Parkinson’s disease the cognitive decline is monitored by considering a simple reach and grasp experiment. This experiment is proposed as non-invasive and cost-efficient test to improve the diagnosis of Parkinson’s disease and monitoring the cognitive decline associated to it. For dystonia a simple hand opening-closing movement is considered as a simple test to detect dystonia, but also the differences between organic and functional dystonia. The idea is to improve the diagnosis of organic and functional dystonia designing an objective, non-invasive and cost-efficient method.

The question will be revisited in the conclusions (chapter 8) using the results available (chapter 5 and chapter 7).

This thesis is organized as follows:

- Chapter 2 describes Evolutionary algorithms in general examining the different program representations that defined the different kinds of algorithm. Cartesian Genetic Programming is the evolutionary algorithm chosen for this work and will be explained in details in section 2.4.

- Chapter 3 gives an overview of Parkinson's disease describing the history of the disease, the cause, main symptoms and also the cognitive impairment associated with it.
- Chapter 4 gives an overview of dystonia describing the history, the different kinds of dystonia with their signs and symptoms, the cause, the diagnosis and the two kind of dystonia considered: functional and organic dystonia.
- Chapter 5 contains a description of the reach and grasp experiment, of the methodology and presents the classification results. The classes considered in the results comprise healthy subjects, Parkinson's disease patients with normal cognition and two subsets of Parkinson's disease patients affected by two different cognitive impairments.
- Chapter 6 describes the wavelet transform which is the method used to pre-process the EMG data collected in the dystonia experiment. Two different wavelet transform are described: the continuous wavelet transform and the discrete wavelet transform. The continuous wavelet transform is described in more detail because we chose to use this approach in this work. Morse wavelet, the analytic wavelet used is described in detail in section 6.1.3.
- Chapter 7 contains the description of the hand opening-closing experiment, of the methodology and presents the classification results. The classes considered are: healthy subjects, organic dystonia patients and functional dystonia patients.
- Chapter 8 contains the conclusions, suggestions for future works and revisits the research question.

## CHAPTER 2: EVOLUTIONARY ALGORITHMS

In this chapter evolutionary algorithms are described. We used these algorithms to evolve classifiers able to distinguish among the different classes considered. The chapter is organised in five main sections. Section 2.1 reports the history of the evolutionary algorithms going through their development over time. Section 2.2 describes evolutionary computation and evolutionary algorithms highlighting their common structure. In Section 2.3 genetic programming (GP), a particular evolutionary algorithm, is described highlighting its representations (section 2.3.1-2.3.2) and its problems (section 2.3.3). Cartesian Genetic programming (CGP), the algorithm chosen in this work is described in section 2.4 evidencing its advantages respect GP. Section 2.5 gives a review of the past medical applications of CGP. Finally the last section 2.6 gives a summary of the chapter with the conclusions and the key points.

### 2.1 HISTORY

There are many different types of evolutionary algorithms (EAs), but all of these algorithms apply the principles of Darwin's theory of evolution by natural selection (Darwin and Bynum 2009). The idea of artificial evolution was introduced the first time by Alan Turing in 1948. Alan Turing wrote the first essay on this topic while he was working on the construction of Automatic Computer Engine (ACE) at the National Physical Laboratory in UK. Unfortunately the essay was dismissed by his employer, who surprisingly was the grandson of Charles Darwin. However, it was later recognized that Turing's essay proposed not only artificial neural networks but all the field of artificial intelligence (Turing 2004). After Turing, the main creators of evolutionary algorithms are Lawrence Fogel, Ingo Rechenberg, Hans-Paul Schwefel and John Henry Holland. Lawrence Fogel in 1966 studied evolving finite state machines used to predict symbol strings (Fogel, Owens and Walsh 1966). Ingo Rechemberg and Hans-Paul Schwefel, between 1971 and 1974 in Germany, gave also a contribution to the development of evolutionary algorithms working in the field of optimization of physical shapes in fluids. They produced good results altering random physical variables, where small changes happened more frequently than larger ones (as in natural selection). This technique was described as a kind of evolutionary algorithm (Eigen 1973, Schwefel 1977). In 1975 in US, John Henry Holland wrote a book about genetic algorithms highlighting the importance of genetic recombination or crossover (Holland 1992).

While Holland in US called his method genetic algorithm, Rechemberg and Schwefel designated their method called evolutionary strategies in Germany. For about 15 years these two areas developed separately. In the early nineties they were unified representing two different representations of the same technology called evolutionary computing.

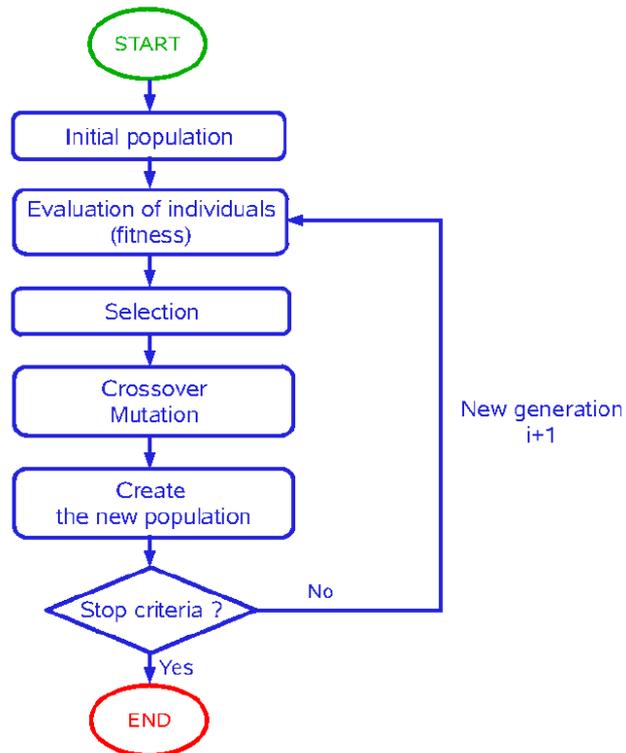
Nowadays evolutionary algorithms are used in many applications, they are popular in solving multi-dimensional problems and also in optimising the design of systems (Onwubolu and Babu 2013, Jamshidi 2003).

## **2.2 EVOLUTIONARY COMPUTATION AND EVOLUTIONARY ALGORITHMS**

Evolutionary computation is the area of study of non-deterministic search algorithms that are based on Darwin's theory evolution of natural selection (Darwin and Bynum 2009). Evolutionary computation represents a family of algorithms for global optimization inspired by biological evolution. An Evolutionary algorithm (EA) is one of the algorithms comprised in the family of Evolutionary computation. It is a generic population-based metaheuristic optimization algorithm. A metaheuristic is a high-level procedure or heuristic (technique designed to solve a problem) with the aim to find an heuristic that could give an acceptable good solution to an optimization problem (Bianchi et al. 2009). So in summary an evolutionary algorithm is an algorithm able to select a heuristic, from an evolving population, which gives us a sufficient solution to an optimization problem. The fitness function in our case is the function to optimize in order to have the best solution. The evolutionary algorithms perform better than the conventional statistical approaches on non-linear highly complex problems (Grosan and Abraham 2008). They are extremely useful also in solving classification and optimization problems with limited or no knowledge of the problem itself.

The basic idea of an evolutionary algorithm is to emulate the natural selection starting from a population of possible solutions of the problem and permitting only the best solutions to survive during the generations. Each solution is called chromosome or phenotype, terminology taken from biology. After the creation of a population, each individual is associated to a value that measures how good the solution is. The function used to evaluate the chromosomes or solutions is called "fitness function". When the fitness function values are calculated for all individuals of the population, the chromosomes with the best values are kept to generate the new generation, becoming "parents" used to create new "children". The generation of new children can be a sexual (crossover) or an asexual process (mutation).

The above process is repeated until a "termination condition" is reached. Termination conditions often include finding a suitable solution and/or a time constraint. The steps of the EAs are resumed in figure 10.



**Figure 10. Representation of the Evolutionary algorithm's steps.**  
Source: [www.canaero.be](http://www.canaero.be)

In an evolutionary algorithm the following concepts have to be defined:

- 1) The initial population
- 2) The fitness function
- 3) The evolutionary strategy
- 4) The reproduction mechanism
- 5) The stop criteria

Usually the initial population is composed by random solutions of the problem. In some cases the initial population can be composed by a set of probable solutions determined by some program.

The fitness function depends on the problem; generally it is necessary to choose a function that measures the grade of “goodness” of a solution.

The evolutionary strategy consists in choosing how many parents have to be selected in each generation and what is the criterion to choose the right parents. For example, we can choose for each generation the two individuals with higher or lower (it depends on the problem) fitness function value as parents of the next generation.

The reproduction mechanism can be mutation (asexual process) or crossover (sexual process). In mutation one parent is involved and the child is simply a

mutated copy of the parent chromosome. In crossover, instead, two parents are involved and the child contains genetic material from both parents with or without mutation.

The stop criteria usually put together a required precision of fitness function value and/or a maximum number of iterations permitted.

## **2.3 GENETIC PROGRAMMING**

In this section an overview of Genetic programming is given. In his book (Miller 2011a) Julian Miller gives an excellent introduction of genetic programming and its different types. The concepts described by Miller are summarized in this section.

Genetic programming (GP) is an evolutionary algorithm with the aim to generate a program which gives the best solution to a determined problem. In 1958 Friedberg for the first time designed a form of GP, describing an algorithm able to evaluate the quality of a computer program (Friedberg 1958, Friedberg, Dunham and North 1959). He designed a mechanism able to make repetitively small random changes in the program and then test them to check possible improvements. Years later in 1980 a kind of GP was used in Smith's PhD thesis (Smith 1980). The first evolved programs having the form of symbolic expression trees were evolved by Cramer in 1985 using TB computer language (Cramer). Despite these first applications, GP become popular only in 1992, when John Koza published his book (Koza 1992).

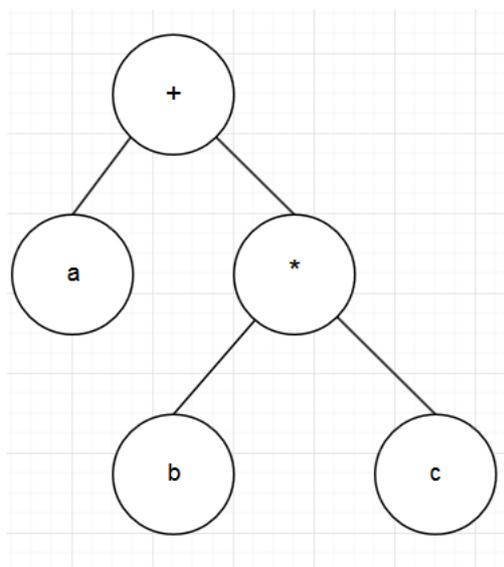
The most challenging part in evolving computer programs is to find the correct representation of the programs, in the way that programs respect specific grammar in order to be compiled. In the two following subsections an overview of different programs representation in GP will be given, while in the last subsection the advantages and the problems of GP will be described.

### **2.3.1 Tree-based representation**

One of the possible representations for programs is the tree representation. It is very easy to represent a program as a tree when it is written in LISP. LISP is one of the oldest high-level computer language and was invented by John McCarthy in 1958 (McCarthy 1960). All the programs in LISP are written as S-expressions, a list of symbols enclosed in parenthesis. Calls of a function are written as the name of the function first followed by its argument. So if we want to call a function *g* with two arguments we should write in LISP (*g arg1 arg2*).

In his book Koza (Koza 1992) described in details the evolution of computer programs written as LISP expression. To understand how easy it is to represent a LISP program as tree data structure, a simple example is considered. The function written in LISP is the following:  $(+ a(* bc))$ , that is equal to the mathematical expression:  $a + (b * c)$ , the correspondent tree representation is illustrated in figure

11. The tree nodes represent the operators while the leaf nodes represent the terminals. To construct the mathematical expression the tree is parsed from left to right.



**Figure 11. Program tree representing the S-expression (+a(\*bc))**

LISP programs, as any program, obey precisely specific syntax and then are highly constrained. The first step is to be able to generate random legal trees (respecting the syntax).

The representation as a tree of the programs simplifies the GP operations. Crossover can be performed easily by exchanging subtrees between parents chromosomes, while mutation is simply the substitution of a subtree with a random one. The size of the chromosome in tree-based GP is variable, because crossover and mutation could result in an offspring of different size.

### **2.3.2 Other representations**

Other than the popular tree representation there are other possible representations in GP that define different kinds of GP.

In linear or machine code GP programs are a constrained linear set of operations and terminals (inputs). They are similar to those written in machine code. In other words programs are codified and represented as bit strings as Banzhaf did originally (Banzhaf 1993). A mechanism of repair has to be generated for the strings because random altered or generated code could result in a syntax violation. Brameier and colleagues defined a recent version of linear genetic programming using variable-length string chromosomes representing simple statements in C programming language (Brameier and Banzhaf 2001, Brameier and Banzhaf 2007). Crossover consists in swapping a whole number of instructions while mutation changes randomly an instruction by replacing the instruction identifier, a variable or a constant, by equivalents from valid ranges.

The grammar-based approaches in GP evolve chromosomes that obey to a specific grammar. A very popular published grammar approach is called grammatical evolution (GE) (O'Neill and Ryan 2001, Ryan, Collins and Neill). In GE, chromosomes are variable-length strings grouped into codons of eight bits. The codon is used as integer value to select a rule from a grammar defined, using a mapping function.

Another less known GP is the PushGP defined using the stack-based computer language Push created by Lee Spector (Spector and Robinson 2002). The interesting aspect of this kind of GP is that is able to support a self-adaptive form of evolutionary computation called auto constructive evolution. The auto constructive evolution permits that the method of crossover and mutation can be evolved in the system rather than be imposed from the beginning.

In the next section Cartesian Genetic Programming will be described in details, but before presenting this technique, we describe its precursor: Cartesian Graph-Based GP. This form of GP is different from the previous ones because graphs are used instead than trees to represent programs. Graph permits having more than one path between any pair of nodes. From a computational point of view this means that, assuming that each node contains a computational function, they permit the reuse of previously calculated computations (contained in sub-graphs). Graph representation is widely used in computer science and engineering (Aho and Ullman 1983, Chartrand, Lesniak and Zhang 2010, Deo 2017), for example artificial neural networks (ANN)(Maind and Wankar 2014) are represented as graphs.

The first person which described a form of Cartesian-Graph based GP was Sushil Louis in 1990 (Louis and Rawlins 1991b, Louis and Rawlins 1991a). He described a binary genotype encoding, a network of digital logic gates, where a gate in each column can be connected to another gate in the previous column. The resulting structure was very similar to a graph.

At the same time, following the studies on neural networks, Poli designed his form of graph-based GP called parallel distributed GP (PDGP) (Poli 1996, Poli 1997). PDGP is a flexible representation able to evolve different kinds of network like tree-like programs, logic networks and neural networks. The flexibility of the approach is given thanks to the introduction of labels associated with the edges of the graph. The labels are dependent on the network to be evolved, for example in neural networks they represent the weights. In order to define a PDGP chromosome there are three things to specify: the set of functions, the set of terminals and the set of links defining how the nodes are connected among each other.

Poli defined the two operations of crossover and mutation for his form of GP. The basic crossover is a generalization of the crossover used in tree-based approach and is called subgraph active-active node (SAAN). In SAAN a random active point is selected from both parents (crossover point), then, from the first parent, the subgraph of active nodes used to compute the output of the crossover points is selected and inserted in the second parent. Poli then defined two versions of mutation: global mutation and link mutation. The global mutation changes

randomly a subgraph in the program while the link mutation changes randomly a connection in the graph selecting a random node and then a random input of the node, changing randomly the link. Details about the SAAN and the two mutation operations are reported in Poli's studies (Poli 1996, Poli 1997).

Cartesian Genetic Programming (CGP), described in the next section, represents programs as graphs like PDGP, but the chromosomes are represented as one-dimensional string and its main operation, the mutation (Poli's link mutation), is done on the chromosome string instead that on the graph itself as PDGP.

### **2.3.3 Advantages and problems**

Genetic programming has a lot of advantages but three main problems:

GP, unlike formal methods, does not need any knowledge about the problem which has to be solved but only a measure to quantify the goodness of a solution. Once initiated GP does not need human interaction. The advantage respect the inductive logic programming is that GP does not perform an exhaustive search inside the problem's solution space; but rather looks for the solution only in the areas which probably contain global optima. These areas are identified through the use of its search history.

However, GP has three main problems:

During its execution, the mechanism of using its search history to explore determined area sometimes fails, because the genetic recombination process does not produce better programs than the existing ones.

The second problem is the bloat problem: the chromosomes that it generates tend to become larger and larger across the generations without any increase in fitness. Typically these chromosomes contain sections of code with inefficient or redundant subexpressions. The evolution of the programs can become very time-consuming and in some cases the program evolved can be bigger than the memory of the computer itself. When the bloat problem occurs it is impossible to find a small and efficient solution and the solutions found could be too hard to understand and not very elegant. Details about the cause of bloat and the possible remedies are reported in the following studies (Langdon and Poli 2002, Poli, Langdon and McPhee 2008, Silva and Costa 2009). In CGP the bloat problem does not occur because genotype has fixed size and also phenotype does not become so large to trigger bloat (Miller and Smith 2006). Turner and Miller also gave an explanation of why CGP does not suffer from the bloat, proposing the two theories of Natural Genetic Drift and Length Bias (Turner and Miller 2014) .

Another problem affecting GP, as other evolutionary computation approaches, is the scalability problem. This means that the increasing problem size results in a high increase of the time and space resources, reaching unmanageable levels.

## 2.4 CARTESIAN GENETIC PROGRAMMING

Cartesian genetic programming was described by Miller's and Thomson's (Miller and Thomson 2000). In this section we describe the main features of CGP, summarizing the popular tutorial (Miller 2013).

CGP differs from the standard tree-based GP, simply for the representation of the programs that are graphs instead of trees. The graph, as in PDGP, permits cycles in a chromosome (cyclic CGP), nodes with arity greater than one and also backward connections between nodes. The only restriction regarding the connections is the fact that the nodes in the same columns cannot be connected to each other.

CGP chromosome is a graph and the shape of the graph depends on the setting of number of columns, number of rows and levels-back. These are denoted by  $nc$ ,  $nr$  and  $l$ , respectively. The maximum number of computational nodes allowed is defined as:  $L_n = nc * nr$ . The parameter levels-back is used to establish how the nodes are connected, determining from which column a node can get its input. For example if this parameter is equal to one then the inputs of a node can come from another node in a column on its immediate left or from a primary input. If instead levels-back is equal to two, a node can be connected to any other node in the immediate left of two columns or to a primary input. Levels-back is set equal to the number of columns when it is allowed the connectivity among any nodes in different columns. In general if there is not a specialist knowledge of the problem, it is convenient to set  $nr=1$  and  $l=nc$  in this case the chromosome is an arbitrary directed graph with a maximum depth (shown in Figure 12).

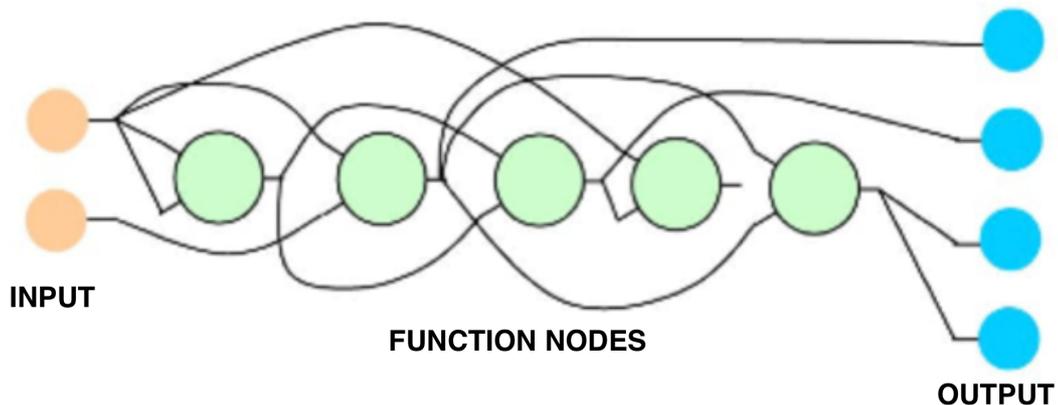


Figure 12. Arbitrary directed graph with a maximum depth.  
Source:(Miller 2013).

The chromosome is a set of nodes of three kinds: input nodes, function nodes and output nodes. The number of inputs, the number of outputs and the set of functions has to be decided.

In CGP each chromosome corresponds to a list of integers called “genotype”. So each chromosome has two representations: a graph representation called phenotype and a list of integers representation called genotype. The genotype is composed by

numbers that are called genes. Each gene can be: a function gene, a connection gene or an output gene. A function gene is a gene representing a function; a connection gene is an input of the function, it can be an input of the chromosome but also an output of a previous node; an output gene instead represents an output of the chromosome.

Figure 13 shows an example of a chromosome's phenotype and genotype.

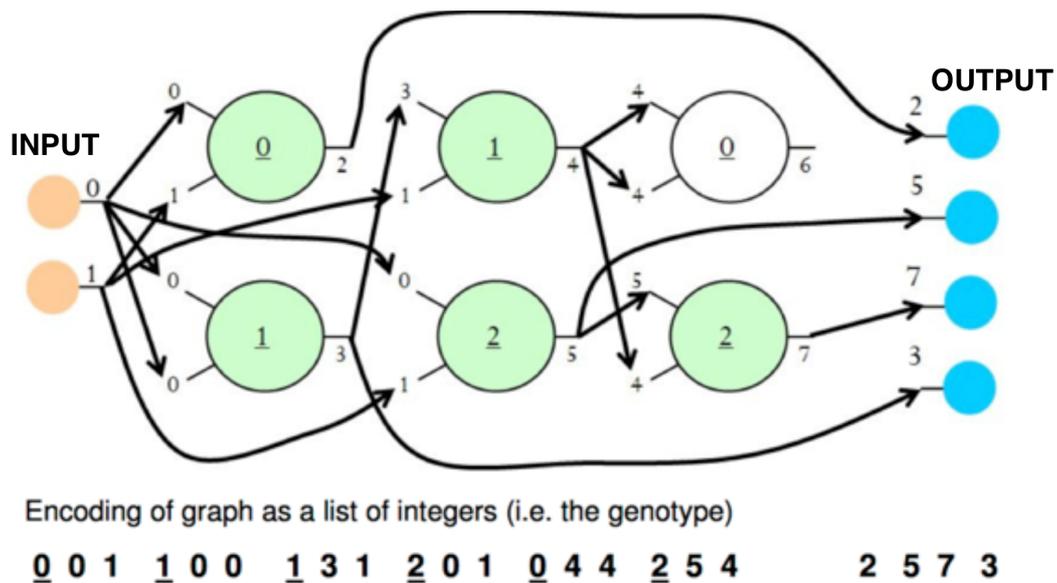


Figure 13. An example of chromosome's phenotype and genotype.  
Source: (Miller 2013).

The meaning of the chromosome phenotype and its associated genotype is easy to understand.

All functions in the chromosome are codified as numbers, in this case: 0 adds data presented to inputs, 1 subtracts data presented to inputs, 2 multiply data presented to inputs and 3 divides data presented to inputs. The genotype in CPG has a fixed length while the phenotype could have a number of nodes from zero to the number of node of genotype (Miller 2011b). The phenotype has zero nodes when all the program outputs are connected to the program inputs (there aren't computational nodes).

Figure 14 helps to understand how the chromosome works, evidencing how the chromosome can be translated in a mathematical expression. This mathematical expression is useful to understand what are the inputs used and how the classifier works. In fact not necessary all the inputs are used to compute the output. Cartesian Genetic Programming as other Evolutionary Algorithms give us the advantage of understanding what are the most important inputs (features in our case) in classifying the different patients. The discover of the important features for each classification case give a better insight in the disorders considered, helping in understanding what are the differences between the patients. This is the main reason of the choice to use Cartesian Genetic Programming to evolve the classifier.

Another advantage of Cartesian Genetic Programming respect the other methods is the easy implementation of the evolved classifier as mathematical expression.

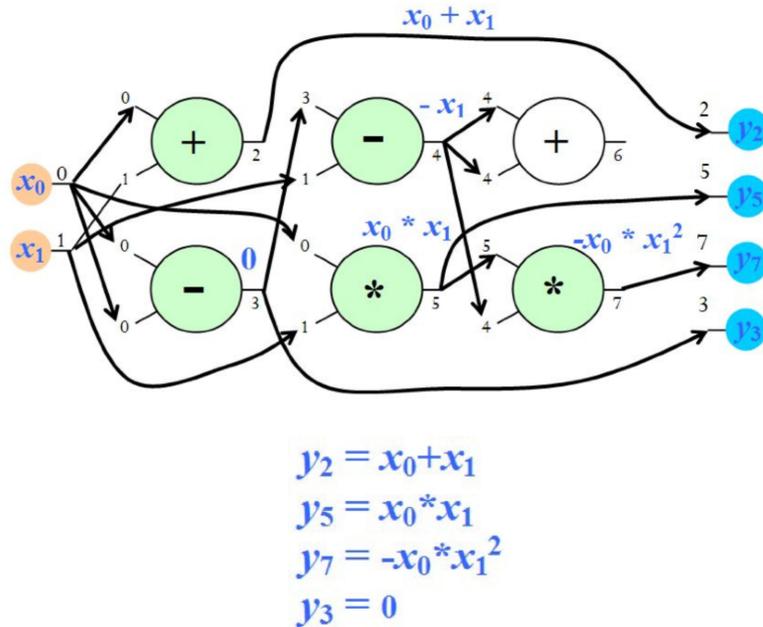


Figure 14. Representation of the chromosome meaning.  
Source:(Miller 2013).

CGP is a kind of evolutionary algorithm so we have to define the features described in section 2.2:

- Initial population: Usually is composed by random chromosomes. Using the inputs and the set of functions pre-defined, the chromosome is created with random connections, random functions (from the functions set) and random outputs.
- Fitness function: Depends on the problem. The function has to measure the goodness of the solution found by the program (chromosome).
- Evolutionary strategy: Often a strategy called (1+4)-ES is used, where in each generation is selected one parent (the fittest chromosome) that is used to create four children (usually by mutation). So each generation has a population of five chromosomes, the parent and its four children.
- Reproduction mechanism: Only mutation is used. The mutation is implemented in a straight-forward way: considering the genotype, each gene has the same probability to mutate (mutation rate). Clearly the mutation has to be “safe”. This means that each mutating gene has to be substituted with a right gene: function gene has to be substituted by another function gene, output gene by another output gene and connection gene by another connection gene.
- Stop criteria: Usually a maximum number of generations are permitted, depending on the shape of the chromosomes.

In Figure 13 the node with the output six does not have any role in the output of the chromosome, then this node is inactive instead all other nodes are active. Clearly there is a many-to-one genotype to phenotype map. So if an inactive node is mutated there is no effect on the phenotype. The mutation of an inactive node is called “silent mutation” because it does not affect the outputs value of the chromosome.

In figure 15 the effect of a silent mutation is depicted. It is clear that this mutation does not change the phenotype but it changes the program connections through subsequent mutational modification. When the genotype is encoded in phenotype the inactive nodes may be ignored to reduce the space of phenotype (Miller 2011b). For example in Figure 13 we could ignore the node six that is inactive. The genes ignored are called “non-coding”. In this case the phenotype and the genotype have the same number of nodes only if all nodes are active.

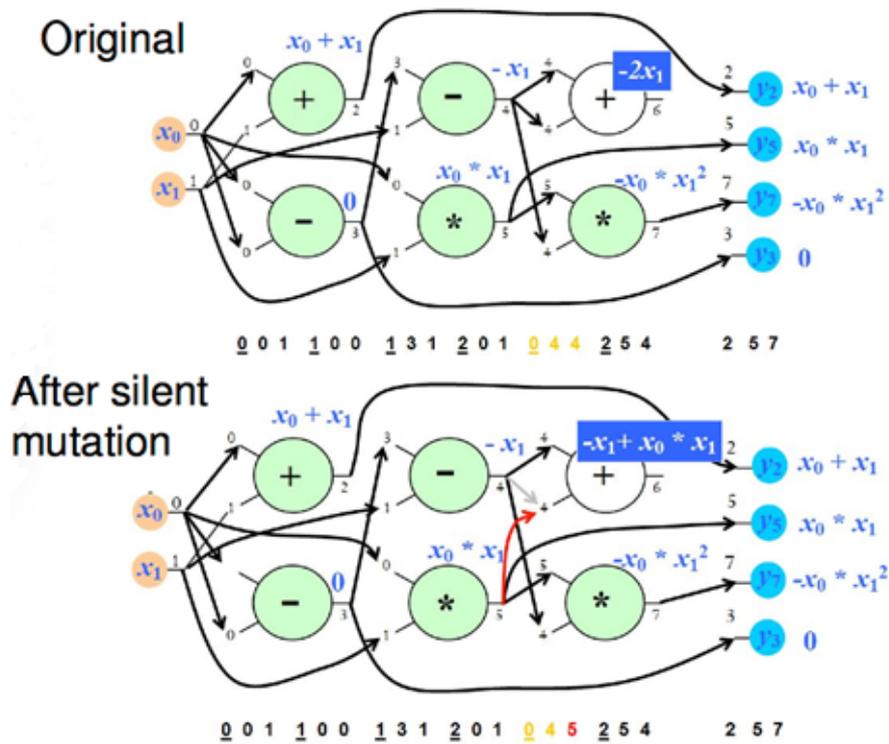


Figure 15. Effect of a silent mutation.  
Source:(Miller 2013).

Instead, when a mutation is done on an active node, it can change a phenotype in a massive way as shown in Figure 16. So with one simple gene mutation the phenotype can change completely with a consequent change of output values. Only mutation is used as operator, because it can give a massive changes in the

chromosome and also because, as demonstrated by Miller (Miller 1999), crossover does not seem to add anything.

However, in some cases, crossover can be very useful: in cases where there are multiple chromosomes with independent fitness assessment (Walker, Miller and Cavill 2006) and also when the floating point representation of CGP is used (Clegg, Walker and Miller 2007).

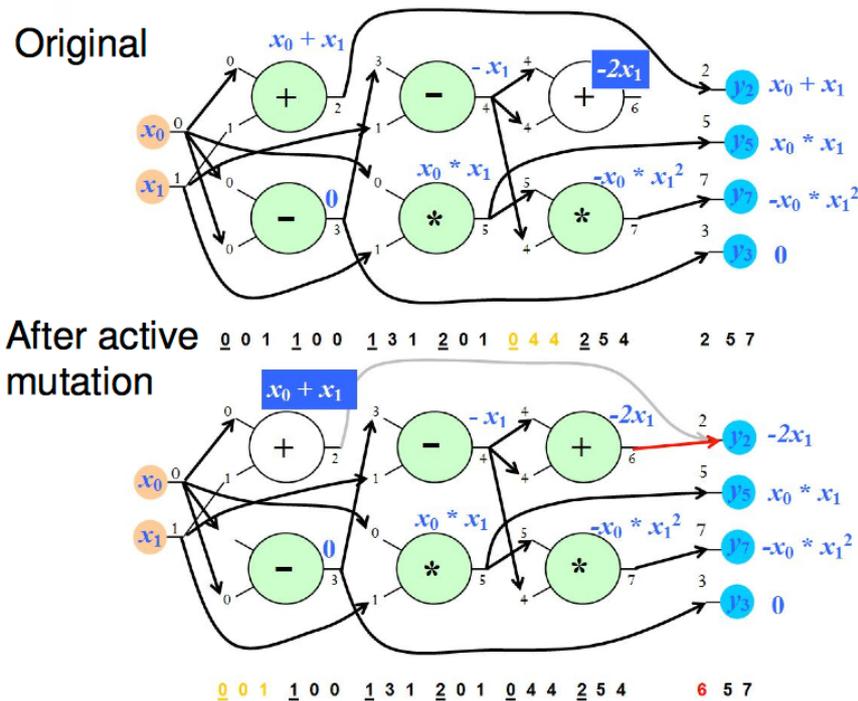


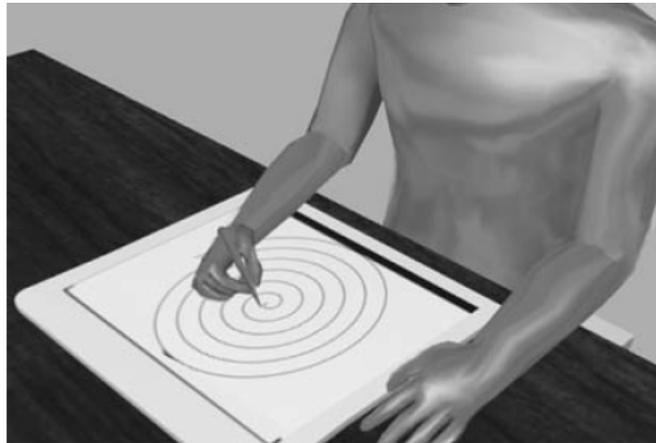
Figure 16. Effect of a single gene mutation on phenotype. Source:(Miller 2013).

## 2.5 PREVIOUS MEDICAL APPLICATIONS

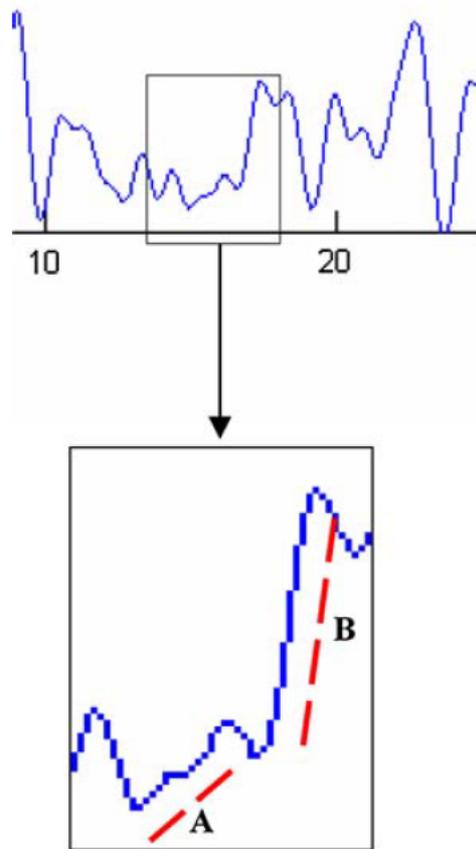
In the past, as mentioned, in my previous study (Picardi et al. 2017), evolutionary algorithms (EAs) were widely used to evolve classifiers in medical context (Zhang and Wong 2008, Paul and Iba 2009, Winkler, Affenzeller and Wagner 2009, Bhowan et al. 2013). Factors which make them very effective for these kinds of problems are: their breadth of search, relatively low sensitivity to initial conditions, and flexibility in terms of representation and evaluation of solutions (Freitas 2009). EAs are extremely useful in all the situations where we do not know precisely the final solution form. The method's breadth of search and the ability to use relatively unconstrained solution representations allow a wide exploration of candidate solutions (Lones et al. 2014). In medical applications usually there is limited understanding on what the solutions look like and then the EAs are extremely useful.

In this work CGP is used taking advantage of the graph representation and of the operations on the genotype. Also, as described in section 2.3.3, the use of CGP permits to avoid the bloat problem.

CGP was used in three previous studies to evolve classifiers able to distinguish between Parkinson's disease patients and healthy subjects with promising results, as revised in our paper (Picardi et al. 2017). In the first study (Smith et al. 2007), a figure-copying task was used, recording pen movements during the experiment. The figure-copying task is represented in figure 17, where the data were collected using a standard digitizing tablet. CGP was applied to find, in the acceleration of the pen, features identifying bradykinesia, a cardinal motor feature of PD (figure 18). In figure 18 a two-part feature of the velocity is shown. This feature evidences a hesitation during the movement, typical of bradykinesia. The limitations of the study are the small dataset and the choice of an experimental threshold based on the dataset.

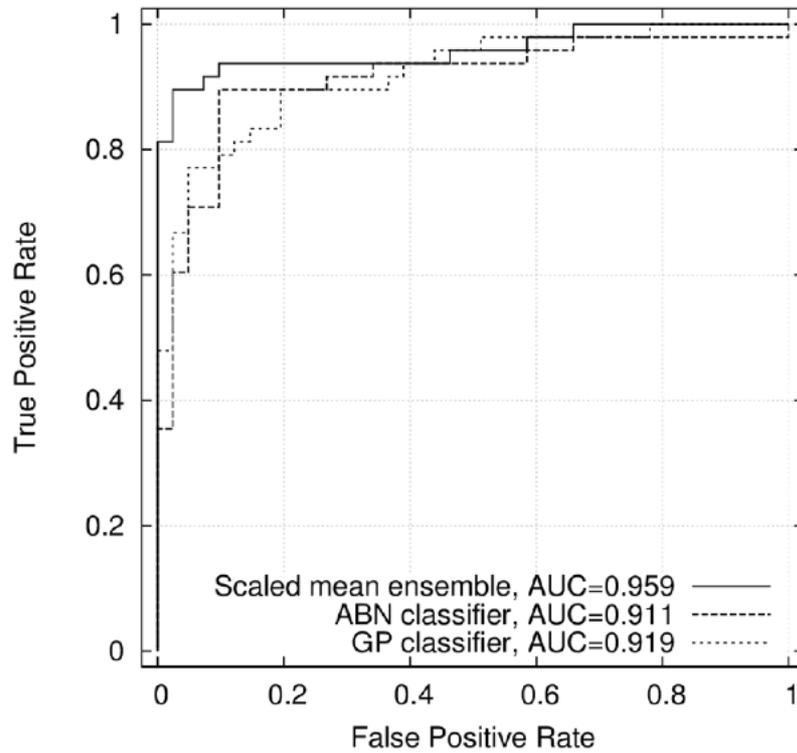


**Figure 17. Figure copying task, the data are measured using a conventional digitizing tablet.**  
Source: (Smith et al. 2007)



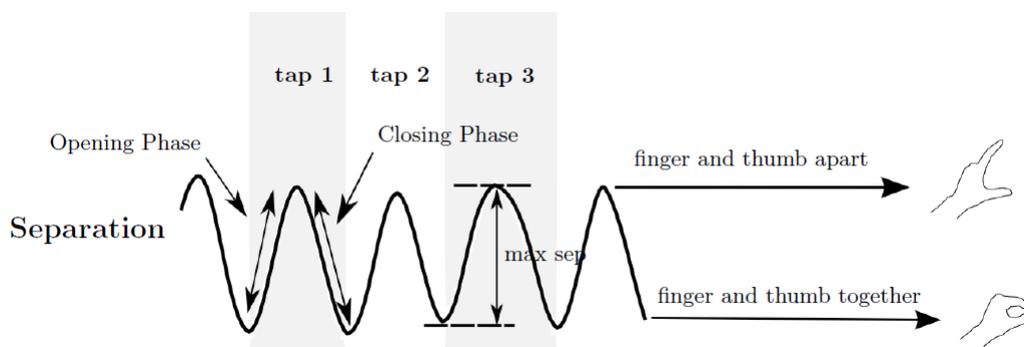
**Figure 18. Example of two-part velocity feature.**  
**This feature evidences bradykinesia. Source:(Smith et al. 2007)**

In the second study (Lones et al. 2014) 49 PD patients and 41 healthy controls performed a finger tapping experiment with each hand separately while the position of the fingers was recorded in real-time. CGP was used to classify PD patients and healthy subjects using as input the acceleration of the fingers, in order to find signs of bradykinesia. In this work two different EAs were used: Artificial biochemical network (inspired by the biochemical networks found in the cell) and a variant of Cartesian Genetic programming. The outputs of the classifiers evolved were combined in ensemble classifiers. The combination was made scaling the output range of each component classifier to the interval  $[0, 1]$  and then using the mean of the classifiers' outputs. The area under ROC curves (Hanley and McNeil 1982) for the ensemble classifier and its component are shown in figure 19. Figure 19 evidences the benefit of combining the results of classifiers evolved by two different EAs, one more able to recognize well characterized local features and the other more sensitive to less evident global features. In this study EAs were able to evolve classifiers which discriminate Parkinson's disease patients and age-matched healthy subjects with an area under ROC curve of 96%.



**Figure 19. ROC curves for the ensemble and its component classifiers.**  
**Source: (Lones et al. 2014).**

The third study (Lacy et al. 2013) also considered finger-tapping. The kinematic data during the experiment were recorded using two electromagnetic sensors one placed on the thumb and the other on the index finger. These data were processed using a Butterworth Low Pass Filter with cut-off frequency 5Hz and then the Euclidean distance was computed and used as measure of separation between the fingers. A trace of digit separation is shown in figure 20 where the single taps, found as intervals between two consecutive local minima, are evidenced. The separation was used as inputs for the Genetic programming in order to evolve the best classifiers. The classifiers evolved in this way were able to achieve an area under roc curve (AUC) (Hanley and McNeil 1982) greater than 0.9.



**Figure 20. Trace of separations values.**  
**Source:(Lacy et al. 2013).**

We used CGP, in this work, to evolve classifiers able to distinguish among PD patients and healthy subjects, PD patients subgroups, dystonia patients and healthy subjects and dystonia patients subgroups. The details on the methodology and the results will be given in chapter 5 and chapter 7. The two movements considered are “reach and grasp” for PD patients and “hand opening-closing” for dystonia patients. In reach and grasp the kinematic data are used while in the hand opening-closing electromyography (EMG) (Gary 2004) data are used.

Reach and grasp is not a repetitive movement as the finger tapping, so the application of CGP is different. It is interesting to evaluate potential of CGP considering a more complex movement (details in chapter 5).

While in hand opening and closing task the aim is to evaluate the potential of CGP in evolving classifiers able to distinguish between dystonia patients and healthy subjects but also between dystonia patients subgroups, taking as input the EMG data (details in chapter 6).

## 2.6 SUMMARY AND CONCLUSIONS

This chapter describes evolutionary computation and in particular evolutionary algorithms. Two particular evolutionary algorithms are described in details: GP and CGP (used in this work). The following list highlights the key points of the chapter:

- Evolutionary computation represents a family of algorithms for global optimization inspired by the biological evolution.
- An Evolutionary Algorithm is one of the algorithms comprised in the family of Evolutionary Computation, able to select a heuristic from an evolving population, which gives us a sufficient solution to an optimization problem.
- The basic idea of an evolutionary algorithm is to emulate the natural selection starting from a population of possible solutions of the problem (chromosomes) and permitting that only the best solutions survive during the generations.
- Genetic programming (GP) is an evolutionary algorithm with the aim to generate a program which gives the best solution to a determined problem.
- The most common program representation in GP is the tree but other representations exist (e.g. Cartesian-Graph based GP).
- GP has the following three main problems: sometimes fails to effectively exploit its search history, bloat problem and scalability problem. CGP solves the bloat problem.
- CGP differs from the standard tree-based GP, simply for the representation of the programs that are graphs instead of trees, permitting cycles in a chromosome (cyclic CGP), nodes with arity greater than one and also backward connections between nodes.

- The chromosome in CGP is a graph easy to convert in a normal string called genotype. The operations on the genotype are easier than on the chromosome itself.
- CGP is used in this work to take advantage of the graph representation and of the operations on the genotype.

In conclusion EAs are chosen to evolve classifiers for their breadth search in the solution space and the relatively low-sensibility to the initial configuration. Therefore EAs are extremely useful when there is poor knowledge on what the solution looks like, situation very common in medical applications. We chose CGP for the graph representation and the consequent derived genotype representation, which makes the mutation operation easy. Also CGP was used in previous similar medical studies with promising results as described in section 2.5. In this thesis we will apply CGP on a more complex movement such “reach and grasp”, not repetitive as the previous movements considered (e.g. finger-tapping), in order to distinguish among subgroups of Parkinson’s disease patients and healthy subjects. In the second part of the work, we will use CGP, considering EMG data as inputs, in order to evolve classifiers able to distinguish among subgroups of dystonia patients and healthy subjects.

## **CHAPTER 3: PARKINSON'S DISEASE**

In this chapter an introduction about Parkinson's disease is given. The aim of the chapter is to explain what Parkinson's disease is and what the main symptoms are. In the subsequent work, detailed in chapter 5, we will evolve classifiers able to distinguish among different subgroups of Parkinson's disease patients and healthy subjects. This chapter gives a better understanding of the disease and the different subgroups of patients considered. It contains four main sections. Section 3.1 reports the history of the disease highlighting the causes and the usual medication with its side effects. Section 3.2 reports the symptoms of the disease describing the four main symptoms: bradykinesia (section 3.2.1), rest tremor (section 3.2.2), rigidity (section 3.2.3) and loss of postural reflexes (section 3.2.4). Section 3.3 describes the possible cognitive problems in Parkinson's disease highlighting the different subgroups measured. Section 3.4, gives a summary of the chapter underlining the key points.

### **3.1 HISTORY OF THE DISEASE**

The first person who described Parkinson's disease was J.Parkinson in 1817 with his essay posthumously published (Parkinson 2002). He described six cases of shaking palsy, describing shaking palsy as: "Involuntary tremulous motion, with lessened muscular power, in parts not in action and even when supported; with a propensity to bend the trunk forwards, and to pass from a walking to a running pace: the senses and intellects being uninjured". The Parkinson's essay describes some symptoms of the disease: resting tremor, abnormal posture and gait, paralysis and diminished muscle strength but also the nature of the disease to progress over the time.

In later years many neurologists contributed to a better knowledge of the disease. One of the most important contribution was given by Jean-Martin Charcot as described by Goetz in his paper (Goetz 2011). Goetz remarks as those Charcot's studies between 1868 and 1881 helped to reach a better comprehension of the disease. In fact as Goetz states: "Jean-Martin Charcot was particularly influential in refining and expanding this early description and in disseminating information internationally about Parkinson's disease. He separated Parkinson's disease from multiple sclerosis and other disorders characterized by tremor, and he recognized cases that later would likely be classified among the Parkinsonism-plus syndromes". Charcot was the first to suggest the use of the term "Parkinson's disease" deciding to reject the term of paralysis agitans or shaking palsy because for him not all the patients presented tremor and weakness. William Gowers, wrote "the Manual of Diseases of the Nervous System" giving his personal experience with 80 patients in the 1980s (Gowers 1898).

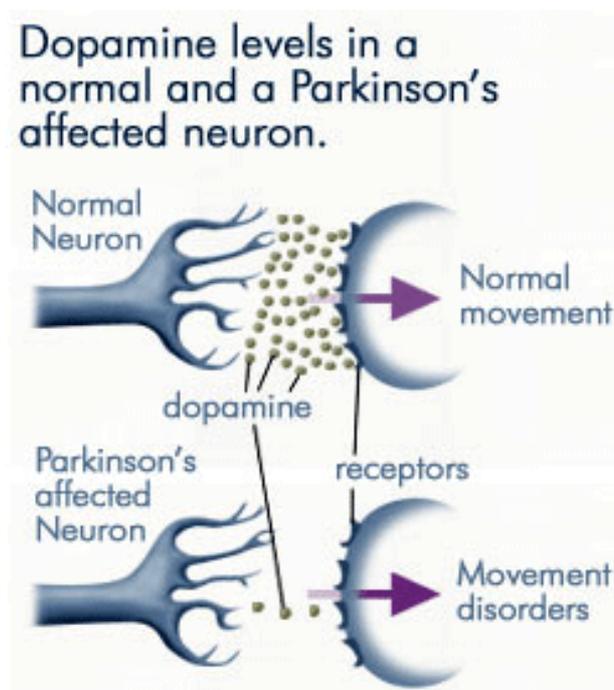
Charcot and William Gowers, with their studies contributed to the early treatments of Parkinson's disease discovering the dopaminergic deficits in the disease. Charchot also helped to distinguish the rigidity from weakness and from bradykinesia (Lees 2007).

In 1912 Frederic Lewy described the microscopic particles of the brain characterising the disease, today called "Lewy bodies"(Lees 2007).

Rolf Hassler's studies in 1938 made clear that the disease affects the substantia nigra and later, in 1950, Arvid Carlsson discovered also the role of the neurotransmitter dopamine in the disease (Fahn 2008).

M. Parents and colleagues (Parent and Parent 2010) remarked the fact that Carlsson and Hornykiewicz contributed to the important discovery that dopamine is a specific chemical marker of nigral neurons that degenerate in Parkinsonism, supporting the view that substantia nigra and Parkinson's disease are closely related.

Following this research, it was clear that Parkinson's disease is a neurodegenerative disease caused by the loss of the neurons producing dopamine, located in the part of the brain called "substantia nigra" (Kim et al. 2003). The loss of dopamine is caused by the degeneration of the dopaminergic neurons that often become abnormal aggregates of protein called "Lewy bodies"(Gibb and Lees 1988). The reduction of dopamine is responsible for the difficulties in the movements; in fact the dopamine facilitates the voluntary movements acting on specific receptors (See figure 21).



**Figure 21. Dopamine levels in normal neurons and in Parkinson's affected neurons.**  
Source: (anti-agingfirewalls.com).

Discovering that the cause of the disease was the deficiency of the dopamine in the brain was very important for the treatment of the disease with the Levodopa synthesized in 1911. Levodopa is the precursor to the neurotransmitters dopamine and is used to increase dopamine concentrations in the treatment of Parkinson's disease. This treatment has been used since the late 1960s to alleviate some motor symptoms of the disease.

With the progression, Parkinson's disease unfortunately becomes less responsive to Levodopa and some side effects such as motor fluctuations and dyskinesia can appear (Liu et al. 2005). The dyskinesia causes involuntary movements that can lead to a very disabling condition.

Liu and colleagues (Liu et al. 2005) examined the effect of the dyskinesia evaluating the involuntary movements during the execution of a spiral-drawing task. They evidenced that Levodopa reduces the effects of the bradykinesia, described in section 3.2.1, but introduces many involuntary movements caused by dyskinesia.

However the side effects of the dyskinesia start after a long-term use of Levodopa (usually at least five years) and probably depend on an increased sensitivity to dopamine, happening when the level of dopamine reaches his peak.

## **3.2 PARKINSON'S DISEASE SYMPTOMS**

J. Jankovic in his paper (Jankovic 2008) described well all the symptoms that differentiate Parkinson's disease (PD) from other parkinsonian disorder. As Jankovic wrote in his work: "Rest tremor, bradykinesia, rigidity and loss of postural reflexes are generally considered the cardinal signs of PD. Other clinical features include secondary motor symptoms (e.g., hypomimia, dysarthria, dysphagia, sialorrhoea, micrographia, shuffling gait, festination, freezing, dystonia, glabellar reflexes), non-motor symptoms (e.g., autonomic dysfunction, cognitive/neurobehavioral abnormalities, sleep disorders and sensory abnormalities such as anosmia, paresthesias and pain)."

In the following sub-sections all the four cardinal signs of PD will be discussed in detail.

### **3.2.1 Bradykinesia**

J. Jankovic described bradykinesia as slowness of tasks during daily life and prolonged reaction times (Cooper et al. 1994), (Giovannoni et al. 1999). As Jankovic says in his work referring to bradykinesia: "this may include difficulties with tasks requiring fine motor control (e.g., buttoning, using utensils). Other manifestations of bradykinesia include loss of spontaneous movements and gesturing, drooling because of impaired swallowing, monotonic and hypophonic dysarthria, loss of facial expression (hypomimia), decreased blinking and reduced arm swing while walking." Bradykinesia is often one of the most evident symptoms

of PD and usually appears before any formal examination.

The assessment of bradykinesia is done through fast, alternate and repetitive movements of the hand (i.e. finger-tapping). During these movements both slowness and decrementing amplitude are evaluated.

Other difficulties related to bradykinesia are: planning, initiating and executing movements, performing sequential and simultaneous tasks. Bradykinesia is strictly correlated to the levels of the dopamine so the effects of bradykinesia decrease with the use of Levodopa. Bradykinesia introduces difficulties to access the motor programmes for PD patients. An external trigger, such as a loud noise, a marching music or a visual cue can help the patients to focus on the action and overcome the difficulties.

In fact many patients, when stimulated, are able to make quick movements (e.g. catch a ball). This was demonstrated by Schettino and colleagues (Schettino et al. 2004). In their experiment a group of healthy subjects and a group of PD patients performed a reach and grasp experiment following a sound. Despite the bradykinesia the PD patients velocity shapes were comparable to the healthy age-matched ones. The results evidenced the severe disruption of internal guidance processes in PD patients. In fact PD patients delayed the execution of the grasp, until visual feedback of their hand was available. The visual feedback appears to be necessary for the subjects in order to modify their hand shape in function of the object to grasp.

### **3.2.2 Rest tremor**

Rest tremor is the most common symptom of the PD. Rest tremor occurs when a muscle is relaxed such as when the hand is resting on a table. Usually tremors are unilateral and are considerable in the distal part of an extremity. In the hand the tremors are like supination-pronation tremors, starting from one hand and then spreading to the other hand. In some cases tremors are also present in the lips, chin, jaw and legs but only in rare cases they affect neck/head and voice. The rest tremor obviously disappears when a PD patient starts an action or sleeps. In some cases there is also an “internal” shaking that is not a visible tremor (Shulman et al. 1996).

Therefore, even if rest tremor is the most common symptom of PD, the occurrence of it is different among the patients and during the disease. In their studies Hughes and colleagues (Hughes et al. 1993) reported that the 69% of patients with PD had rest tremor at onset of the disease however, 75% of patients manifested rest tremor during the progression of the disease. Rest tremor disappears during the disease in 9% of PD patients. Martin and colleagues in their study evidenced that a small percentage of patients (11%) never had tremor (Martin et al. 1973), but another study demonstrated through the use of autopsy that 100% of patients had tremor during the disease (Rajput, Rozdilsky and Rajput 1991).

Another possible tremor that can affect the patients is postural tremor. Postural tremor is a kind of essential tremor (Deuschl and Elble 2009), it happens when a

patient tries to hold a part of his body fighting the gravity force (e.g. holds an arm in front of you). This kind of tremor is more prominent and disabling than rest tremor and usually is present when the subject assumes an outstretched horizontal position (Jankovic, Schwartz and Ondo 1999, Jankovic 2002). Postural tremor has the same frequency of the more common rest tremor and for this it is responsive to dopaminergic therapy. It may be one of the first symptoms of Parkinson's disease.

### **3.2.3 Rigidity**

Rigidity affects the ability of the muscle to relax, resulting in tensing muscles unable to relax properly. Most PD patients with the progression of the disease develop some degree of rigidity or stiffness of limbs. The rigidity limits the movement ability and is caused by involuntary and uncontrolled tensing of muscles (Lance, Schwab and Peterson 1963). Aches or pains could be also present in the affected muscles. Rigidity could be associated to hyperactive stretch reflexes in the muscles concerned (Rushworth 1960).

Increased resistance causes the rigidity and often the "cogwheel" phenomenon is present. When the "cogwheel" phenomenon is present, a movement is not continuous but is subdivided in a small fragmented movement remembering the cogwheel of the clock (Lance et al. 1963).

Rigidity of the neck and trunk (axial rigidity) could cause abnormal axial postures (e.g., anterocollis, scoliosis). Postural deformities are caused by the rigidity and can result in flexed neck, posture, knees and elbows. However, flexed posture generally occurs in the final stages of the disease (Jankovic 2008).

### **3.2.4 Loss of postural reflexes**

Postural Reflexes are automatic movements that control the equilibration and permit to fight the gravity force when the subject stands up and moves.

Loss of postural reflexes causes a postural instability which usually occurs after the onset of other clinical features. The postural instability is the major cause of falls and so increases the risks of bones fractures (Williams, Watt and Lees 2006). There are some risk factors to the development of postural instability. People with diabetes and PD have an higher risk to develop postural instability than patients having only PD (Kotagal et al. 2013). Cognitive impairment, described in subsection 3.3, seems to increase also the risk to develop postural instability (Owan et al. 2015). Additionally another study shows that PD patients with depression and apathy have a major risk to develop postural instability (Hassan et al. 2014). At the onset of PD it is possible fight this symptom using the PD medications, but with the progression of the disease they become less effective. Some studies shows that physical therapy can be useful to fight the postural instability improving balance and preventing falls, but it is still unclear if the effects of the therapy are permanent or only temporary (Klamroth et al. 2016).

Some of the symptoms described above can cause the freezing phenomenon that when associated with gait is another cause of falls. As Jankovic wrote in his work: “Freezing, also referred to motor blocks, is a form of akinesia (loss of movement) and is one of the most disabling symptoms of PD.”

A good way to describe freezing is imagine that during walking the feet get “glued to the ground”. When this happens, patients cannot move forward also for several seconds, they feel their feet “frozen” or stuck, but at the same time they are able to move the top half of the body (Parkinson's UK). Freezing can manifest when patients start to walk or try to turn around but it is not limited to walking. Sometimes patients freeze while they eat, speak or perform other repetitive movements such as writing (Parkinson's UK). Rigidity, bradykinesia, postural instability and longer duration of the disease represent the increase the risk of developing freezing, while tremor at the beginning of the disease seems to decrease the risk (Jankovic 2008).

Whereas the above symptoms are the most known, few people know that the PD patients have also many problems in the coordination between limbs. Van den Berg and colleagues in their study (van den Berg et al. 2000) investigated the problem of coordination between limbs in PD. Their experiment consists in rhythmic forearm movements performed by 11 PD patients and 11 healthy subjects. The movements were executed in three different ways (considering both phase and anti-phase movements): comfortable amplitude in-phase, anti-phase and single-arm mode at pacing frequencies ranging from 0.5 to 3 Hz. Results evidenced coordination problems in PD patients. For both in-phase and anti-phase modes the PD group performances were significantly worse respect to the control group (healthy subjects). Also the PD group showed a greater variability of relative phase than the control group.

### **3.3 COGNITION IN PARKINSON’S DISEASE**

It is very difficult to define what the cognitive impairment is. In general cognitive impairment causes difficulties in remembering, learning new things, concentrating and making everyday decisions.

In this work two kinds of cognitive impairment associated with PD are considered: Parkinson’s disease mild cognitive impairment (PD-MCI) formally defined in 2012 (Litvan et al. 2012) and Parkinson’s disease dementia (PDD) (Docherty and Burn 2010).

Patients with mild cognitive impairment notice some changes in cognitive functions starting to forget things and having problem in concentrating. Anyway these changes do not severely affect daily activities, permitting the subject to live independently. The situation is different for the patients with dementia that present a severe impairment and lose the ability to manage daily living activities resulting in losing the ability to live independently (Cognitive Impairment Parkinson’s

Disease Foundation).

Cognitive impairment is very common in PD. Several studies demonstrate that approximately 50 % PD patients develop dementia in ten years from diagnosis and 80 % in 20 years time (Williams-Gray et al. 2013, Perez et al. 2012). PD-MCI is recognised to be present in a percentage between 35% and 42.5% of people at the time of the diagnosis of the disease (Broeders et al. 2013, Yarnall et al. 2014).

Janvin in his study shows that in some cases PD-MCI could lead to PDD (Janvin et al. 2006), in particular the evolution of PD-MCI in PDD seems to depend on which cognitive domain is affected. Usually when PD-MCI affects memory, language and visuospatial domains there is a greater risk that this can evolve in PDD (Williams-Gray et al. 2007).

An early detection of the cognitive impairment is very important to maximise the benefit of any possible therapy.

There is a grown interest in trying to identify potential biomarker to diagnose the cognitive impairment as soon as possible with the aim of maximising the medication effects.

Previous studies investigated the correlation between gait and cognition (Amboni et al. 2012, Amboni, Barone and Hausdorff 2013). In the first study the gait of patients with and without mild cognitive impairment are compared showing that the patients with impairment have reduced step-length, reduced swing time and dynamic stability. The second study also shows that cognitive impairment is connected to postural instability. This is probably due to the fact that the impairment affects the visuospatial domain.

In the first part of the work described in this thesis, we will distinguish among three different subgroups of PD patients and healthy subjects using a number of features. The groups measured are: PD-NC (Parkinson's disease patient with normal cognition), PD-MCI, PDD and a group of healthy subjects called control group. The idea is that this classification can identify potential biomarkers relative to the impairment, considering a reach and grasp task. The reach and grasp experiment is done also with eyes closed (memory-guided reach way) to find a better evidence of impairment effects.

The classifier evolved to distinguish between the patients with different cognitive impairment can be used to monitor the impairment itself. The patient can be monitored periodically with the test evidencing if there are any improvements in the impairment by the examination of the classification results. This monitoring is particularly useful when a new drug is given, testing periodically the patient to see if the classification result changes and the cognitive impairment consequently improves. This monitoring can be very useful to judge the effectiveness of new drugs helping in the research of a cure for these disorders.

### 3.4 SUMMARY AND CONCLUSIONS

This chapter contains an introduction to Parkinson's disease describing the causes, the symptoms, the medications used and the possible cognitive decline. In the following list the key points are highlighted:

- Parkinson's disease (PD) is a neurodegenerative disease caused by the loss of the neurons producing dopamine, located in the part of the brain called "substantia nigra".
- The dopaminergic neurons degenerates in abnormal aggregates of protein called "Lewy bodies"
- The reduction of dopamine is responsible for the difficulties in the movements; in fact the dopamine facilitates the voluntary movements acting on specific receptors.
- Levodopa is used to increase dopamine concentrations in the treatment of Parkinson's disease, but it could lead to dyskinesia causing involuntary movements leading to a disabled condition.
- The main symptoms of Parkinson's disease are: bradykinesia, rest tremor, rigidity and loss of postural reflexes.
- Two kinds of cognitive impairment associated with PD are considered: Parkinson's disease mild cognitive impairment (PD-MCI) and Parkinson's disease dementia (PDD).
- PD-MCI patients notice some changes in cognitive functions (e.g. forgetting things), but these changes are not so bad to influence the subject independence.
- PDD patients present a severe impairment, which causes the loss of the ability to manage daily living activities leading to inability of living independently.

In conclusion PD is a neurodegenerative disease, which affects the brain and causes difficulties in performing voluntary movements. The evolution of PD in time could cause cognition problems important to detect in order to manage them with the proper drugs.

## **CHAPTER 4: DYSTONIA**

In this chapter an introduction about dystonia syndrome is given. In the second part of the work described in this thesis we use CGP (section 2.4) to evolve classifiers able to distinguish among subgroups of dystonia patients and healthy subjects. The chapter is useful to understand dystonia itself and the different kinds of dystonia along with their symptoms and signs. Also functional and organic dystonia, the two subgroups considered, are described highlighting their differences.

The chapter is organised in five main sections. Section 4.1 reports the history of the disease defining the disease itself. Section 4.2 reports the symptoms and signs of dystonia relative to its different kinds. Section 4.3 describes the causes and the way to diagnose the syndrome dividing it in two big categories primary (section 4.3.1) and secondary (section 4.3.2) dystonia. A different kind of dystonia called functional dystonia is reported in section 4.4 highlighting its differences respect to the common organic dystonia. A summary of the chapter with key points and conclusion is reported in section 4.5.

### **4.1 HISTORY OF THE DISEASE**

In 1713 Bernardino Ramazzini for the first time described dystonia syndrome (Ramazzini 2001). In his book he described diseases related to different workers. In particular he noticed that clerks and in general all the workers who sit and write for a long time could develop in time “failure of power in the right hand”. This failure for Ramazzini was linked to “the incessant movement of the hand and always in the same direction” and “the incessant driving of the pen over the paper” that causes “the continuous and almost tonic strain on the muscles and tendons”. This description represents a task-specific dystonia, a focal dystonia that affects one or more muscles in a part of the body causing involuntary muscle contractions and abnormal postures. The task-specific dystonia is present usually during particular motor activities involving highly skilled, repetitive movements (Torres-Russotto and Perlmutter 2008). Writer’s cramp is diagnosed as task-specific dystonia and Ramazzini in his book for the first time described it. In 1814 Samuel Solly recognised this phenomenon in writers calling it the scrivener’s palsy (Pearce 2005).

All historical reports related motor abnormalities to the overuse of muscles.

In 1911 Hermann Oppenheim (1858–1919) for the first time introduced the term dystonia describing the characterizing spasms of the syndrome (Oppenheim 1911). In 1975 there was the first international conference on dystonia, in which also the task-specific dystonia known as the writer’s cramp, in addition to the severe forms, was recognised as kind of dystonia (International Parkinson and Movement Disorder Society (MDS) 2013).

In the following years all the different dystonia syndromes were identified and the

clinical complexity of the disease was recognised.

At this stage it was clear that dystonia is a neurological movement disorder where excessive and involuntary muscle contractions result in repetitive movements and/or abnormal postures (The National Institute of Neurological Disorders and Stroke 2018). Neychev and colleagues (Neychev et al. 2008) in their study, highlighted that dystonia can be caused by a disruption of a motor network involving both basal ganglia (part of the brain involved in controlling voluntary movements) and cerebellum (part of the brain playing an important role in motor control of the movements). The symptoms can involve a restricted group of muscles and with time spreading to the adjacent muscles (Balint and Bhatia 2014).

There is a growing interest in discovering how psychological factors can affect dystonia. In their work R.E. Newby and colleagues (Newby et al. 2017) reported a complete review of dystonia's history remarking how two main intellectuals Jean Martin Charcot (1825–1893) and Sigmund Freud's (1856–1939) had a long lasting effect on thinking about dystonia. Charcot separated the “organic” disorder, which could be related to structural changes in the nervous system, from the “functional” disorders that could not. A lot of dystonia cases over the next two decades were identified as “névroses”: his term for conditions without an identifiable neuroanatomical cause. Sigmund Freud's developed theories about how psychic distress and past experiences could be transformed in physical symptoms. These theories were applied to dystonia making it unclear if the disease had neurological or psychiatric cause. More details about how the psychological factors could cause symptoms and generate a kind of dystonia called “functional dystonia” will be given in section 4.4.

## **4.2 KINDS OF DYSTONIA SIGNS AND SYMPTOMS**

Symptoms of dystonia can be different accordingly to the kind of dystonia. All the symptoms described in this section are present on the website of the dystonia society (The Dystonia Society 2018b).

One of the possible symptoms is an abnormal blinking or twitching of the eye that clears up quickly. Often there is the feeling that the eye is tired or dry and the eyelid is heavy. These symptoms are usually present in the eye dystonia (Blepharospasm) characterized by an increased blinking and involuntary closure of eyelids. Sensitivity to bright lights can also be present in eye dystonia.

The Cervical Dystonia (Spasmodic Torticollis) is instead characterized by painful contractions of the muscles in the neck. This can bring the feeling that the head/neck is pulling to one side, backwards or forward and also painful muscle spasms or stiffness in the neck. The subjects could develop abnormal posture of the head/neck and difficulty to turn the head to one side with relative easy turning to the other side.

The hand dystonia (otherwise known as Writer's or Musician's Cramp), described in the previous section is characterized by the twisting or curling up of the hands

during a certain activity such as writing or playing an instrument. The fingers can move independently assuming unusual position during the activity and the activity undertaken could result in pain. Usually all the symptoms of the hand dystonia disappear when stopping the particular movement, such as writing.

In voice dystonia (laryngeal dystonia or spasmodic dysphonia) there are voice problems. The voice can be strained with tremors; the subject makes big effort to speech resulting in a strangled voice with the feeling of choking on words. Sometimes it is difficult to speak in noisy environment because the voice results to be breathy, whispering or very quiet.

There are symptoms that could be present in all different kinds of dystonia such as pain, cramping and muscle spasms due to involuntary muscle movements. Sometimes tremor is also present (uncontrollable shaking oscillation of a part of the body) and is called dystonic tremor. Dystonic tremor could affect the same part of the body affected by dystonia but also other parts. Usually the oscillations are irregular and the tremor is asymmetric (affecting only one arm for example instead of both). This tremor could be temporarily removed by a sensory trick such as a light touch on the chin for head tremor.

Severe attacks of dystonia, called “dystonic storms” could be present in subjects where the dystonia affects a great part of the body. During these attacks the subject is conscious but usually unable to communicate because the muscles of the face and vocal cords are involved. This status is called dystonicus and its causes are unknown, it is believed that some change in medication can trigger it. The medications are used to reverse this status and in some rare case, where medications do not work, deep brain stimulation is used (Apetauerova et al. 2010).

As described in the symptoms above many kinds of dystonia exist. The different kinds of dystonia are classified in different way. One of the ways is accordingly to the affected area of the body (Brain Foundation 2018): Focal dystonia affects only one part of the body such as neck (Cervical dystonia); Segmental dystonia affects adjacent body parts; Multifocal dystonia involves more than one unrelated body parts; Hemidystonia affects one arm and one leg on the same side of the body and Generalized dystonia affects more or less all of the body.

The age onset of dystonia is strictly related to the body area affected and to the spreading of the diseases (Geyer and Bressman 2006). As Geyer and colleagues wrote in their study, usually early-onset in a childhood dystonia starts in a leg or arm and rarely from neck, vocal cords or other face muscles; while on the contrary late-onset in adults starts in the neck or in general cranial muscles and rarely starts in a limb such as leg. Also the spreading of the disease depends on the age onset with the early-onset dystonia usually spreading from one limb to another and becoming generalised in 50% of the cases (Geyer and Bressman 2006). Late-onset dystonia usually do not spread all over the body remaining focal or segmental (Jayne, Lees and Stern 1984, Friedman and Fahn 1986).

Like many movement disorders all the symptoms described above could be worsened by fatigue and emotional stress, and at the same time could be improved

by sleep and relaxation (Geyer and Bressman 2006). Usually to treat dystonia drugs are used to reduce muscle spasms and pain, but also Levodopa in some particular kinds of dystonia.

## **4.3 CAUSES AND DIAGNOSIS OF DYSTONIA**

The cause of dystonia cannot be identified in all patients; also it is very important in order to ensure a personalised effective treatment. Dystonia can be divided in two basic categories: primary or idiopathic dystonia and secondary or symptomatic dystonia (Geyer and Bressman 2006). Geyer and colleagues gave an excellent description of the two categories and how to diagnose dystonia in both cases. Their description will be summarized in the following subsections.

### **4.3.1 Primary dystonia**

Primary dystonia presents only signs of dystonia itself. This means that only abnormalities related to proper dystonia are detected. If other signs of parkinsonism, seizures, dementia, ataxia, spasticity or others are found it is more likely to be secondary dystonia. The only symptom that can be present in primary dystonia, along with the proper dystonia symptoms, is tremor or myoclonus. However usually myoclonus or tremor movements are less predominant than dystonia, if this is not the case dystonia is more likely to be secondary dystonia.

Focal or segmental dystonia are the most frequent primary dystonia with late-onset. Only 10% of patients have generalised primary dystonia with childhood or adolescence age onset (Nutt et al. 1988).

Although many patients do not have a family history of dystonia a genetic cause is identified or at least suspected for many primary dystonia (Bressman et al. 1989, Defazio et al. 1993, Stojanović, Cvetković and Kostic 1995). Usually dystonia is diagnosed as primary when the cause is genetic or not identifiable.

The diagnosis of primary dystonia first involves genetic test to check if dystonia is one of the genetic identified dystonia, then if the test is negative other tests are undertaken to exclude secondary dystonia. One of the tests performed to exclude secondary dystonia and other pathologies is the MRI of the brain, useful to exclude structural lesion and signal abnormalities suggestive of other syndromes such as metabolic syndromes.

### **4.3.2 Secondary dystonia**

Secondary dystonia is diagnosed when other abnormalities are identified other than the proper dystonia ones. This kind of dystonia usually can be induced by many factors providing insult to the brain (Dystonia Medical Research Foundation 2018):

- Spinal cord, head and peripheral injury (trauma)
- Drugs such as levodopa in treatment of Parkinson's disease (levodopa-induced dystonia), or dopamine receptors blocking agents drugs
- Other pathologies affecting the brain like cerebral palsy, cerebral infections, stroke, brain tumor etc.
- Neurological or metabolic diseases such as Huntington's disease and Wilson's disease

Secondary dystonia has many subcategories; one of these is the dystonia-plus syndrome. In these kinds of disorders dystonia is present with other neurological abnormalities, but, as in primary dystonia, there are no signs of brain degeneration. Dopa-responsive dystonia is one of the dystonia-plus syndromes. This kind of dystonia is easy treatable, usually with levodopa is possible to restore all or almost all normal motor functions over time (Segawa et al. 1976).

Another subcategory of secondary dystonia contains many inherited disorders in which it is possible to find a sign of brain degeneration. These disorders cause abnormal movements due to the disruption of the basal ganglia function or to the interference with dopamine synthesis. Usually examining the neuroimaging it is possible to notice some alterations and dystonia appears to be less prominent than other conditions.

As reported above also other degrading factors can cause secondary dystonia in healthy subjects. Brain insults could be caused by: drugs such as block dopamine receptors, infections, trauma and others.

However secondary dystonia is also associated to other movement diseases. Usually it is associated with Parkinson's disease, often causing a painful foot dystonia when levodopa concentrations are low, but it could affect other body areas.

The diagnosis for a secondary dystonia may start when the assessment for the primary one is negative or when the clinical features suggest a secondary dystonia.

Although there are several factors that cause secondary dystonia as described above, usually an analysis of the age-onset, family history and presence of other features such spasticity should be sufficient to diagnose a secondary dystonia. MRI of the brain can be also useful to determine abnormalities.

#### **4.4 FUNCTIONAL AND ORGANIC DYSTONIA**

In the previous sections we described different kinds of dystonia that are all part of organic dystonia. Organic dystonia, as described previously, could be primary or secondary and in both of these cases there are appropriate tests to diagnose it.

Functional dystonia instead is a condition where there are some specific symptoms of dystonia, but the tests establishing the cause of the symptoms are negative (The

Dystonia Society 2018a). The symptoms are similar to those of organic dystonia, but they are due to psychological factors rather than a brain disease. In some cases functional dystonia and its symptoms are triggered by a psychological trauma.

The diagnosis of functional dystonia is not easy. Usually it is diagnosed by a neurologist with experience in this kind of dystonia. Ganos and colleagues (Ganos et al. 2014) in their study highlighted the fact that the differentiation between organic and functional (psychogenic) dystonia is not clear. This fact leads to a misdiagnosis of both organic and functional dystonia, causing a bad management of the condition wrongly diagnosed. In fact for functional dystonia a physical and/or occupational therapy is preferred, other possible therapies are cognitive behaviour therapy and/or oral medication such as anti-depressant. As Ganos and colleagues (Ganos et al. 2014) explain some features characterizing functional dystonia could help its diagnosis. Usually in functional dystonia symptoms appear suddenly and are often preceded by physical and/or emotional event. There could also be variability in progression and duration with influence of life-related events on remissions and recurrences. Other functional phenomena such as functional tremor could be present and a suggestibility to the placebo or an atypical response to medications could be observed. Finally, functional dystonia may coexist with other organic movement disorders or neurological illness affecting the same side of the body.

An early diagnosis of functional dystonia could be crucial to address the cause and reduce frustration in patients. It is also important, when it is possible, to discover the psychological cause of the dystonia itself, because usually treatment of the cause leads to disappearing of the symptoms. The psychological cause is not ever easy to discover because not all functional dystonia forms are accompanied by psychological symptoms such as anxiety and depression.

However, if the symptoms are caused by psychological factors they are real and appear exactly as the organic dystonia symptoms. In order to demonstrate that functional dystonia can also cause real physical problems, Schrag and colleagues (Schrag et al.) examined neuroimaging to find some differences between functional and organic dystonia. Participants comprised 6 functional and 5 organic dystonia right leg patients along with 6 matched healthy control subjects. They were measured during rest, fixed posturing of the right leg and during paced ankle-movements. Both functional and organic patients showed abnormal activity in prefrontal cortex rejecting the hypothesis that these abnormalities are a marker for the functional disorders. The findings of this study are really important, because, as Dr. Schrag added in a statement, "opens up a way for researchers to learn how psychological factors can, by changing brain function, lead to physical problems".

The second aim of this work is to distinguish among healthy subjects, functional and organic dystonia patients. This is done by measuring the activity of two specific muscles during hand opening-closing task. Different features are extracted and the classifiers are evolved using the evolutionary algorithms described in Chapter 2 and two other well-known machine learning algorithms: Artificial Neural Networks (ANN) (Maind and Wankar 2014) and Support Vector Machine (SVM)

(Durgesh and Lekha 2010). The details of the experiment, the method and all the results are given in Chapter 7. The aim is defining a methodology useful to appreciate the differences between healthy subjects and dystonia patients but also between organic and functional dystonia patients. The classification of the two subgroups of dystonia patients is very challenging, but also really interesting. In fact this classification could be used to understand better the differences between the groups, helping the doctors in difficult diagnosis.

## **4.5 SUMMARY AND CONCLUSIONS**

The aim of this chapter is to give an introduction about dystonia. The following list highlights the key points of the chapter:

- Dystonia is a neurological movement disorder syndrome where excessive and involuntary muscle contractions result in repetitive movements and/or abnormal postures.
- There are different kinds of dystonia (e.g. cervical dystonia, hand dystonia), each kind affects specific muscles and has different symptoms.
- There are symptoms that could be present in all different kinds of dystonia such as pain, cramping, tremor and muscle spasms due to involuntary muscle movements.
- Usually to treat dystonia drugs are used to reduce muscle spasms and pain and also Levodopa in some particular kind of dystonia.
- Dystonia can be divided in two main categories: primary or idiopathic dystonia and secondary or symptomatic dystonia.
- Primary dystonia presents only signs of dystonia itself. This means that only abnormalities related to proper dystonia are detected.
- Secondary dystonia is diagnosed when other abnormalities are identified other than the proper dystonia ones. This kind of dystonia usually can be induced by many factors providing damage to the brain.
- Primary and secondary dystonia are both part of organic dystonia.
- In functional dystonia the symptoms are similar to those of organic dystonia, but they are due to psychological factors rather than to a brain disease. In some cases functional dystonia and its symptoms are triggered by a psychological trauma.
- It is important but also challenging to distinguish between organic and functional dystonia in order to design the best therapy for the subject.

In summary dystonia is a neurological disorder, which affects muscles causing involuntary movements, cramps and pain. Functional dystonia is due to psychological factors but presents all the normal symptoms.

# CHAPTER 5: REACH AND GRASP EXPERIMENT

In this chapter the first part of the thesis work is detailed. We considered a reach and grasp experiment in three different visual conditions. Three different subgroups of Parkinson's disease patients PD-NC, PD-MCI, PDD (section 3.3) and a group of healthy subjects were measured recording all the kinematic data. The aim is distinguishing among the different subclasses using some features extracted from the data and given as inputs of the evolutionary algorithms (Chapter 2).

The chapter is divided in four main sections. Section 5.1 describes reach and grasp movement reviewing the previous studies. Section 5.2 contains the methodology describing: the subjects recruited (section 5.2.1), the equipment used (section 5.2.2), the experiment (section 5.2.3), the reaching data (5.2.4), the features extracted from the reaching data (5.2.5) and the grasping data (5.2.5) explaining why they are not used in this study. Section 5.3 reports all the results and it is divided in three subsections: in the first subsection (5.3.1) all the data belonging to all the different visual conditions are analysed to determine if one of the conditions is more discriminating, in the second subsection (5.3.2) all the results are reported and the third subsection (5.3.3) shows a diagram of a classifier evolved with the derived mathematical expression. Section 5.4 reports the conclusions and the suggestions for future work.

## 5.1 REACH AND GRASP MOVEMENT

Reach and grasp task can be defined as the movement done to reach an object and grasp it. This movement, also called prehension task, is a simple movement composed by two components (Marteniuk et al. 1990, Jeannerod 1986):

- **Reach or transport component:** where the brain commands the responsible upper limb muscles to transport the hand from the start point to the target. The usual kinematic parameters of this phase are: movement time, peak (maximum) velocity, peak acceleration, peak deceleration, times to reach all the peaks from the movement onset (time to peak velocity etc.).
- **Grasp or manipulation component:** where the brain commands the muscles responsible for the movements of the fingers to start the grasp phase (open and close hand). The usual kinematic parameters of this phase are: peak aperture (maximum distance between index and thumb), time to peak aperture from the movement onset, manipulation time (time from the reaching onset to the first sign of hand opening saw as the first separation sign between index and thumb).

Prehension task is very suitable for diagnosing Parkinson's disease because the patients present great difficulties to perform manipulative task. These difficulties are evident in the differences of reach and grasp parameters (movement time, peak velocity, peak aperture, peak acceleration etc.) between PD patients and healthy subjects.

The delayed starting of the manipulation component in PD patients was evidenced by Castiello and colleagues (Castiello, Stelmach and Lieberman 1993). In their study, they evaluated the prehension performance of PD patients and noticed that the patients started to open the hand (manipulation component) only after approximately 80-90 ms from the transport onset. The healthy subjects instead started the manipulation component approximately 30 ms after the transport onset. This resulted in a prolonged movement time for PD patients.

An interesting study explores the effect in response of PD patients to a double unexpected perturbation in position and size of the target object (Castiello et al. 1999). With the double perturbation the PD patients reached the peak acceleration on average 50 ms earlier than the normal case. The response to the perturbations in the grasp component appeared 500 ms after that of the transport component. On the other hand, the response of the healthy subjects to the perturbations appeared approximately simultaneously in the transport and in the grasp phase prolonging only the deceleration phase. The results supported the view that PD patients have difficulties in performing coordinated actions, assuming the basal ganglia to be part of the circuit that modulates the degree to which the components are coordinated according to output requirements.

Also other studies suggested that basal ganglia played an important role in the execution of coordinated task as reach and grasp (Teulings et al. 1997, Alberts 1997, Gentilucci and Negrotti 1999). So with a basal ganglia dysfunction, like that caused by Parkinson's disease, there is a difficulty in performing coordination movements resulting in a disruption in both temporal and spatial domain.

A study examined this disruption effects for PD patients (Isenberg and Conrad 1994). The authors considered simple pointing movements with different movement distances and speeds comparing the performances of the PD patients with the healthy subjects. The results reported more segmented reaching paths for PD patients than for healthy subjects. The irregularity of the reaching paths for PD patients become worse as the speed increased, probably because with a greater movement speed the time for single components synchronization is less. In addition PD patients spent more time in the vertical plane (raising their arm) than in horizontal plane (reaching the target). Healthy subjects on the contrary performed movements using both planes simultaneously and resulting in the normal curvilinear reach path.

The visual conditions (presence or absence of visual feedback) influence the reach and grasp parameters of PD patients, as reported in the following studies.

The lack of visual feedback seems to destroy the coupling between reach and grasp (Castiello, Bennett and Mucignat 1993). Jeannerod demonstrated that when the

visual feedback is removed the hand opening depends on the object size (greater for bigger objects than for smaller ones), but it is greater with respect to the full vision case (Jeannerod 1984). The movement time was prolonged when the visual feedback was absent. The same findings were confirmed in another study (Jakobson and Goodale 1991).

The increased hand opening could be a way to avoid the collision between fingers and object, but also a way to compensate the reach error, opening the hand to grasp a bigger object (Fukui and Inui 2013).

Schettino et al. investigated the effect of visual feedback for PD patients in reach and grasp (Schettino, Adamovich and Poizner 2003). In their experiment, subjects had to reach and grasp three different shaped objects (a rectangular cube, a convex shaped block and a concave shaped block) under three different visual conditions (full vision, in darkness with the exception of a visually illuminated target and without any visual feedback). Movement time was prolonged when visual feedback was removed both completely and partially (only object is illuminated). Also the time to peak aperture occurred earlier in the same cases supporting the theory that a margin of error is employed in the absence of visual feedback (the hand opens wider and sooner in these conditions). These recent findings then confirmed the previous ones in literature (Jeannerod 1984, Jakobson and Goodale 1991).

In summary the kinematic reach and grasp parameters affected by the visual feedback seem to be movement time and hand opening. Movement time increases when visual feedback is absent (Jakobson and Goodale 1991, Jeannerod 1984, Schettino et al. 2003) or when the room is dark and only the object is illuminated (Schettino et al. 2003). Peak aperture (maximum aperture of the hand) is greater and occurs earlier in absence of visual feedback (Jakobson and Goodale 1991, Jeannerod 1984, Schettino et al. 2003).

In this study the reach and grasp is examined considering different visual condition and different speeds. The idea is to try to classify the PD patients and the healthy subjects using some representative features of the reach and grasp, which as the previous studies demonstrated, are affected by the basal ganglia dysfunction present in PD. These features, which are computed using kinematic parameters, are used also to distinguish among the different subgroups of patients: PD-NC, PD-MCI and PDD (see section 3.3). Some of them were computed in an old study on apraxia to distinguish between apraxia patients and healthy subjects but also among different subgroups of apraxia patients (Caselli et al. 1999). The apraxia patients were slower than the healthy subjects and the slowness is reflected in movement time, reaction time but also in all kinematic transport and manipulation parameters. Minor differences were found among apraxia patients subgroups.

The set of features of Caselli's study was not so powerful to distinguish among the classes considered. We decided to extend this set with other features relative to the angular velocity, acceleration and to the jerk (time derivative of the acceleration).

Alberts et al. (Alberts et al. 2000) computed angular velocity and jerk finding some interesting differences between PD patients and healthy subjects. They investigated

the reach paths of PD patients considering different accuracy constraints (bigger or smaller objects). The jerk score, a feature computed from the jerk was used to estimate the smoothness of the reaching paths. The smoothness was also evaluated examining the regularity of the angular velocity profile. They confirmed the previous findings in which the PD patients have less smoother reaching paths with respect to the healthy subjects (Isenberg and Conrad 1994). Also reaching paths were more segmented when accuracy increased. In addition they highlighted the differences in kinematic patterns of the transport and the grasp component, evidencing that PD patients were slower than healthy subjects.

In this study 25 features are computed using the previous studies (Caselli et al. 1999, Alberts et al. 2000) and used as input to the classifier, evolved with the technique explained in Chapter 2, with the aim to distinguish among all the classes considered (PD, PD-NC, PD-MCI, PDD and controls). The methodology with the details about the experiment and the features computed will be described in the following sections.

## **5.2 METHODOLOGY**

### **5.2.1 Subjects recruited**

This thesis is based on the study ‘A novel diagnostic device for the objective diagnosis of Parkinson’s disease with and without dementia’, which received National Regional Ethics Service approval (reference code 10/H1308/5) and local Research and Development approval from Leeds Teaching Hospitals NHS Trust (LTHT) (reference code UI10/9232). The subjects were measured in Leeds, UK (Cosgrove 2016).

Patients were recruited from Dr Jamieson’s and Dr Alty’s current consultant caseloads and also from other neurology clinics at Leeds Teaching Hospitals NHS Trust. They are divided in three classes according to the cognition: PD-NC, PD-MCI and PDD and attended specialised clinics and day hospital that retain the capacity for informed consent. The clinical researcher Dr Cosgrove was responsible to measure the subjects ensuring that all the patients read the protocol and signed the informed consent.

Fifty-eight PD patients and 29 healthy subjects called controls were assessed between February 2014 and October 2014. Controls were recruited from the spouses and friends of PD patients who attended clinic.

The Movement Disorders Society – Unified Parkinson’s disease Rating Scale (MDS-UPDRS) – Part 3 was used to assess the motor symptoms of all participants. This is a validated scale that assesses the motor signs of PD (Goetz et al. 2008).

The cognition of the patient was evaluated using two different scales: Montreal Cognitive Assessment (MoCA), a global test to assess the cognitive function (Nasreddine et al. 2005), and Clinical Dementia Rating Scale (CDR), a semi-structured interview designed to assess different grade of dementia from normal

cognition to severe dementia (Morris 1997).

MoCA score was used to separate the patients with normal cognition (PD-NC with MoCA  $\geq 26$ ) from those with some cognitive impairment (PD-CI with MoCA  $< 26$ ) (Cosgrove 2016). The global CDR score was used instead to separate the PD-CI group; a score of 0 or 0.5 was categorised as PD-MCI and a score of  $\geq 1$  was categorised as PDD (Cosgrove 2016).

In total the fifty-eight patients were divided in: 22 PD-NC, 23 PD-MCI and 10 PDD. Three patients were impossible to classify according the two scales because they had limit value scores; they are then excluded from the categories. They were considered in the class containing all PD patients. In table 1 all the classes considered with the demographical details are summarized.

	<b>Controls</b>	<b>PD-NC</b>	<b>PD-MCI</b>	<b>PDD</b>
<b>Age, years</b>	63.8(7.9, 50-75)	66.5 (9.4, 44-84)	70.0 (8.0, 47-85)	72.6 (5.3, 64-83)
<b>Gender, M:F</b>	4:15	16:6	14:9	6:4
<b>Handedness, R:L</b>	15:4	20:2	20:3	8:2
<b>Duration disease, years</b>	-	5.1 (3.7, 0.5-15)	5.7 (4.0, 0.5-15)	10.5 (6.4, 1.0-20)
<b>Number subjects</b>	29	22	23	10

**Table 1. Summary of the class considered with demographic details (Standard deviation, range).**

Part of the table is taken from (Cosgrove et al. 2016).

## 5.2.2 Equipment used

The experiment performed is a simple reach and grasp experiment. Measurements of the reach and grasp motor tasks were made by using a commercially available computer data glove 5DT Data Glove 5 Ultra manufactured by Fifth Dimension Technologies (California, USA) with an incorporated electromagnetic tracking system manufactured by Polhemus Inc. (Vermont, USA).

In Figure 22 the glove is shown, it is made of Lycra, comfortable to wear with exposed finger tips to allow fine motor actions such as manipulation of objects. Lycra gloves were made to fit all hand sizes.

The 5DT Data Glove 5 Ultra has five sensors, one on each finger to measure its flexion (used to measure the manipulation component). The sensor computes an average finger flexion value for all fingers (Fifth Dimension Technologies, 2014). Similar technology in a glove was used in previous reach and grasp studies (Schettino et al. 2003, Schettino et al. 2004).



**Figure 22. 5DT Data Glove 5 Ultra.**

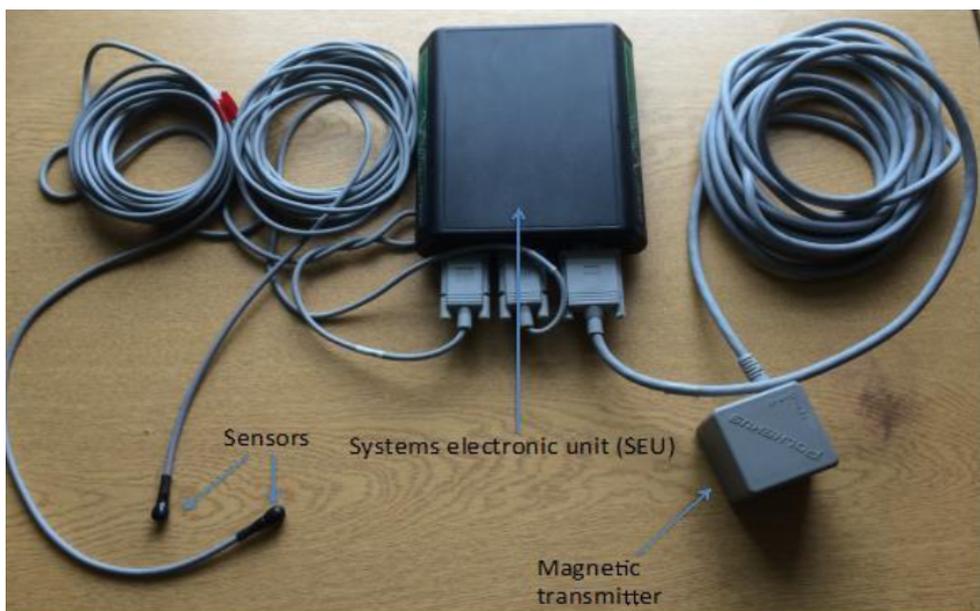
**It allows real-time tracking of hand movements whilst maintaining the flexibility and convenience of an existing clinical test environment. Source: [www.5dt.com](http://www.5dt.com).**

The computerized data glove incorporated an Electromagnetic (EM) sensor placed on the wrist (Figure 23), in order to track the position and orientation of it during the experiment (used to measure the transport component).

The EM sensor is part of the Polhemus Patriot an EM tracking device composed by a system electronic unit (SEU), two sensors (one of them placed on the wrist) and a magnetic transmitter (Figure 24). The sensor placed on the wrist permits location with six degrees of freedom: three positional coordinates (x, y and z) and three orientation coordinates (roll, pitch and yaw). The position of the wrist is computed as distance between the sensor on the wrist and the magnetic transmitter that was placed five centimetres behind the object to reach.



**Figure 23. EM sensor fixed on the wrist under the Velcro strap of the glove.  
Picture produced by Dr Jeremy Cosgrove (Cosgrove 2016).**



**Figure 24. The EM tracking device Polhemus Patriot.  
Picture produced by Dr Jeremy Cosgrove (Cosgrove 2016).**

### **5.2.3 Experiment description**

The reach and grasp experiment required the subjects to reach an 8 cm diameter cylinder (such as baker) placed on a table 30 cm in front of them, grasp it, lift it and put it back on the table. The experimental setup is illustrated in Figure 25. The subject was positioned in front of the table, sat on a chair, with the hands semi-pronated rested on the table. The little fingers rested on the table in the positions corresponding to 2 and 3 marks depicted in Figure 25.

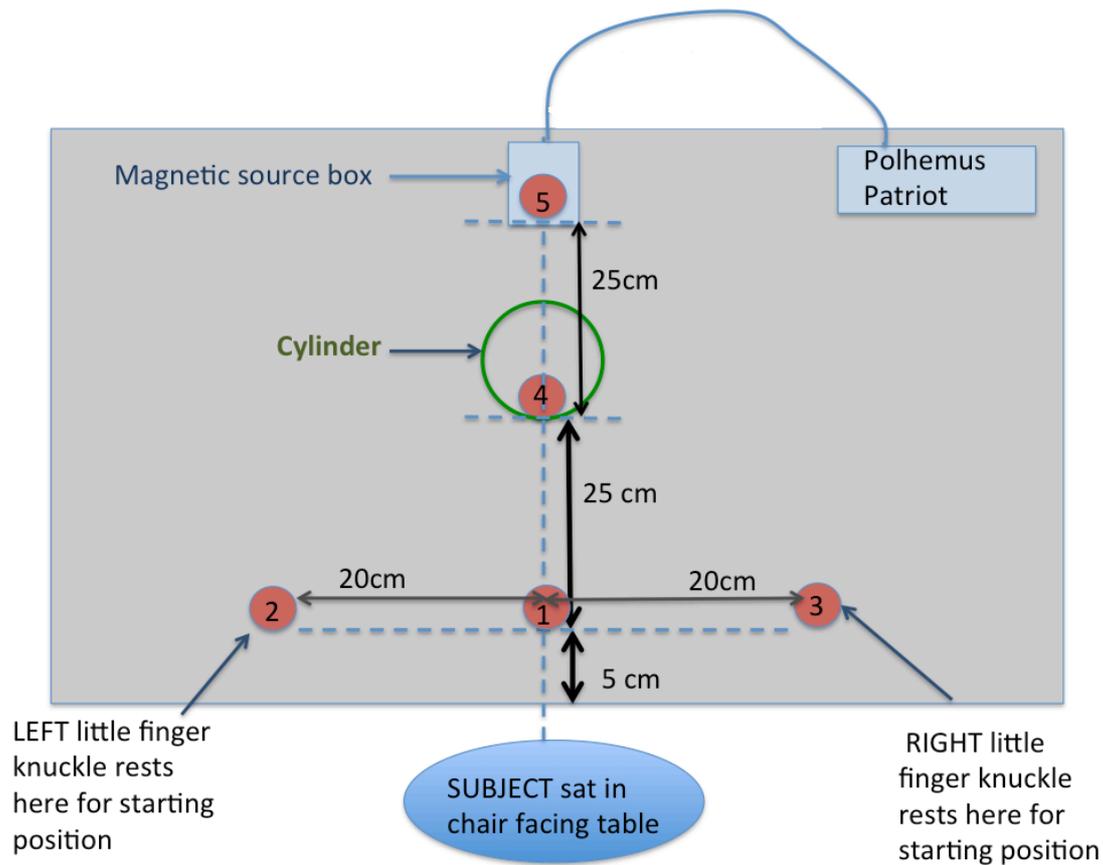


Figure 25. Experiment setup.  
Picture produced by Dr Jeremy Cosgrove.

Each subject did the experiment five times with their dominant hand and five times with their non-dominant hand under four different conditions:

- **Self-guided reach1 (NAT):** instructions were to reach and grasp the object as they would be at home grasping a beaker (natural speed) after a sound clue.
- **Visually cued reach (VIS):** participants had to reach and grasp the object after it light up (red light). The room was dark as much as possible and there was also a simultaneous sound clue as in the other cases.
- **Self-guided reach2 (MAX):** instructions were to reach and grasp the objects as fast as they could after the sound clue.
- **Memory guided reach (MEM):** participants were asked to close their eyes before the start of the task. They had to reach and grasp the cylinder keeping their eyes closed. Only when the cylinder was put back on the table were they allowed to open their eyes. The signal to start the experiment was the audio clue as in the previous cases.

In summary each subjects did the task ten times (five for each hands) for each conditions (forty repetitions in total).

## 5.2.4 Reaching data

The patriot sensor, placed on the wrist, recorded the values of six different spatial coordinates: three positional coordinates (x,y,z) and three rotational coordinates (roll, pitch, yaw). The six degrees of freedom are shown in figure 26 and figure 27.

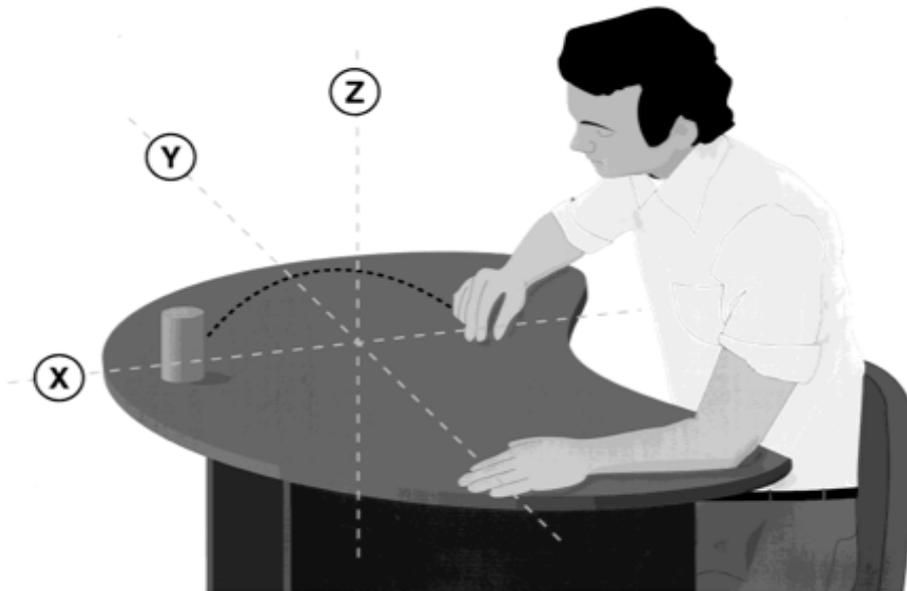


Figure 26. Positional space coordinates (x,y,z).  
Source: (Caselli et al. 1999)

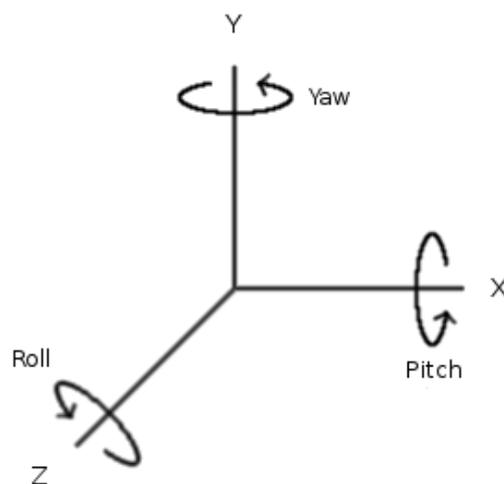


Figure 27. Rotational coordinates (roll,pitch,yaw).  
Source: (Pegasus 3D mesh 2017)

All the data discussed in this subsection and the features that will be discussed in the next subsection were used for the results published in 2017 (Picardi et al. 2017).

Using the reaching data recorded in the experiment the following measures were computed:

- **Distance:** the positional coordinates  $(x,y,z)$  are used to compute the position, the velocity and the acceleration of the wrist during the experiment. The position is computed as the Euclidean distance between the sensor on the wrist and the EM magnetic source placed behind the cylinder:

$$\text{dist}(t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2} \quad (5.1)$$

- **Velocity:** each velocity component  $(v_x, v_y, v_z)$  is computed differentiating each positional component  $(x,y,z)$  considering  $h = 1/60$  seconds as sampling time. Then the velocity along the x direction at the time  $t_1$  is:

$$v_x(t_1) = \frac{x(t_1 + h) - x(t_1)}{h} \quad (5.2)$$

The velocity is then computed as a modulus using all its components:

$$\text{vel}(t) = \sqrt{v_x(t)^2 + v_y(t)^2 + v_z(t)^2} \quad (5.3)$$

Only the modulus of the velocity is considered because the direction is known.

- **Acceleration:** is computed differentiating the velocity described above with the same sampling time  $h$ . So the acceleration at the time  $t_1$  is:

$$\text{acc}(t_1) = \frac{\text{vel}(t_1 + h) - \text{vel}(t_1)}{h} \quad (5.4)$$

In this case the sign of the acceleration is important because in order to distinguish between acceleration ( $\text{acc}(t)>0$ ) and deceleration ( $\text{acc}(t)<0$ ).

- **Jerk:** it is the first time derivative of the acceleration. Jerk is computed differentiating the acceleration using the sampling time  $h$ . Jerk value at the time  $t_1$  is then:

$$\text{jerk}(t_1) = \frac{\text{acc}(t_1 + h) - \text{acc}(t_1)}{h} \quad (5.5)$$

- **Angular velocity:** the rotational coordinates roll  $\varphi$ , pitch  $\theta$  and yaw  $\psi$ , are used to construct the rotation matrix describing the pose of the wrist with the three rotations around ZYX. The Aerospace MATLAB Toolbox (Aereospace MATLAB Toolbox 2017) is used to compute the matrix. The skew-symmetric matrix  $S(t)$  (Siciliano 2009) is computed from the rotation matrix  $R(t)$  with the following formula:

$$S(t) = \dot{R}(t)R^T(t) \quad (5.6)$$

The matrix  $S(t)$  is computed each 1/60 seconds and contains the components of the angular velocity vector  $\omega = [\omega_x, \omega_y, \omega_z]^T$  that are its symmetric elements, with respect to the main diagonal, in the form (Siciliano 2009):

$$S = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \quad (5.7)$$

The angular velocity magnitude  $\omega$  is then computed from its components:

$$\omega(t) = \sqrt{\omega_x(t)^2 + \omega_y(t)^2 + \omega_z(t)^2} \quad (5.8)$$

- **Angular acceleration:** is computed by differentiating the angular velocity, described above, using the same sampling time  $h = 1/60$  seconds. Then the angular acceleration at the time  $t_1$  is:

$$\text{ang}_{\text{acc}}(t_1) = \frac{\omega(t_1 + h) - \omega(t_1)}{h} \quad (5.9)$$

A Savitzky-Golay filter is used to smooth all velocities and accelerations (Luo, Ying and Bai 2005). The smoothing is done fitting successive sub-sets (windows) of adjacent data points with a low-degree polynomial using the linear least square method (Luo et al. 2005). A window size of 35 and degree of polynomial 5 was

chosen throughout experimentation. The experimentation was done with an iterative method incrementing gradually the window size and the polynomial degree. The window size of 35 and a polynomial degree equal to five gave the best results smoothing the signal without changing the shape in order to avoid loss of information.

### 5.2.5 Features extracted

The measures described in the previous subsection 5.2.4 were used to compute the different features that are described in detail in this subsection. The features extracted are related only to the reaching data because the use of grasping data was more difficult than we thought at the beginning of the work. Details about the grasp data will be given in the next subsection 5.2.6.

Figure 28 shows the distance profile of one trial made by a control (healthy subject): on x-axis the time in seconds is reported, while on y-axis the distance (from the starting point) in millimetres is reported. The distance decreases when the subject starts to move until a minimum is achieved when the subject reaches the object (end of the reach phase). The distance then increases when the subject lifts the cylinder and decreases at the end when the subject leaves the cylinder on the table (end of the lift phase). The following points are highlighted:

- **Stimulus:** the dotted line marks the time at which the cue occurs. The cue is a sound or a light (for the visually cued task) used as a signal for the subject to start the movement.
- **Start of the movement:** instant at which the subject starts to move. This point is identified as the point in which the distance between the sensor and the source starts to decrease.
- **End of the movement:** instant at which the subject reaches the object. This point is identified as the first minimum of the distance between the sensor and the source.

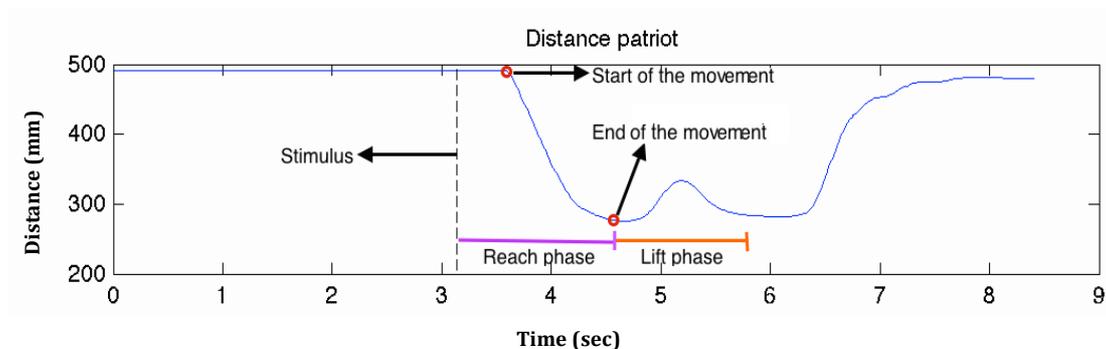
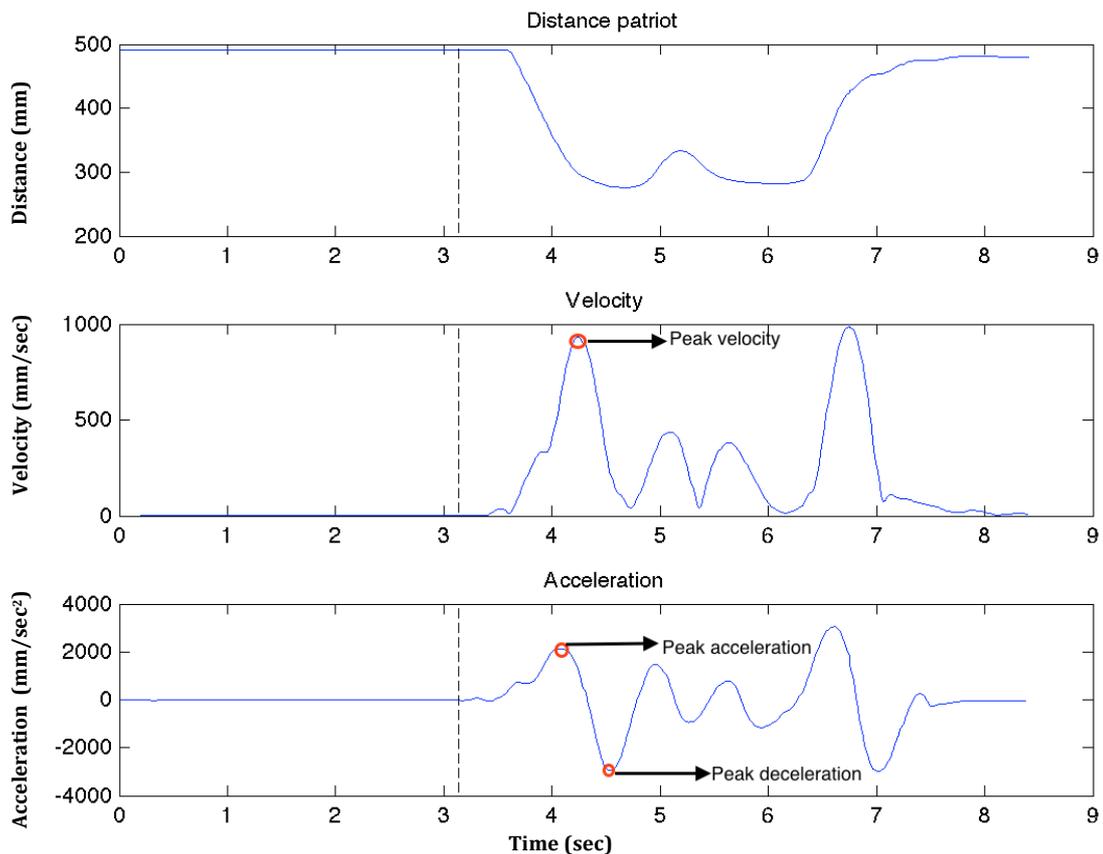


Figure 28: Distance profile relative to a trial made by a healthy subject.  
Source: (Picardi et al. 2017).

Figure 29 considers the same movement examined in Figure 28 with velocity (on x-axis time is measured in seconds, on y-axis velocity is measured in millimetres/second) and acceleration (on x-axis time is measured in seconds, on y-axis acceleration is measured in millimetres/second) derived from it. In the graph the following points are highlighted:

- **Peak velocity:** The maximum of the velocity identified in the interval of time between the start of the movement and the end of the movement.
- **Peak acceleration:** The maximum of the acceleration identified in the interval of time between the start of the movement and the end of the movement.
- **Peak deceleration:** The minimum of the acceleration identified in the interval of time between the start of the movement and the end of the movement.



**Figure 29. Distance (upper), velocity (middle) and acceleration (bottom) profiles relative to the same trial of Figure 28. Peak velocity, acceleration and deceleration are evidenced. Source: (Picardi et al. 2017).**

Using the points detailed in Figure 28 and Figure 29, the following features are computed:

1. **Movement time (MT):** The time taken between the start of the movement and the end of the movement.
2. **Reaction time as a percentage of total movement time (RT%):** The time taken between the cue and the start of the movement expressed as a percentage of the total movement time.
3. **Peak velocity (PV):** The value of the velocity at its peak.
4. **Time to peak velocity (TPV):** The amount of time taken from the start of the movement to the point where the velocity peak occurs.
5. **Time to peak velocity as a percentage of total movement time (TPV%):** The amount of time taken from the start of the movement to the point where the velocity peak occurs, expressed as a percentage of the total movement time.
6. **Peak angular velocity (PAV):** The maximum of the angular velocity found in the interval of time between start of the movement and end of the movement.
7. **Time to peak angular velocity as a percentage of total movement time (TPAV%):** The amount of time taken from the start of the movement to the point where the angular velocity peak occurs, expressed as a percentage of the total movement time.
8. **Time to peak angular velocity (TPAV):** The amount of time taken from the start of the movement to the point where the angular velocity peak occurs.
9. **Peak acceleration (PA):** The value of the acceleration peak.
10. **Time to peak acceleration as a percentage of total movement time (TPA%):** The amount of time taken from the start of the movement to the point where the acceleration peak occurs, expressed as a percentage of the total movement time.
11. **Time to peak acceleration (TPA):** The amount of time taken from the start of the movement to the point where the acceleration peak occurs.
12. **Peak angular acceleration (PAA):** The maximum of the angular acceleration found in the interval of time between start of the movement and end of the movement.
13. **Time to peak angular acceleration as a percentage of total movement time (TPAA%):** The amount of time taken from the start of the movement to the point where the angular acceleration peak occurs, expressed as a percentage of the total movement time.
14. **Time to peak angular acceleration (TPAA):** The amount of time taken from the start of the movement to the point where the angular acceleration peak occurs.
15. **Peak deceleration (PED):** The value of the deceleration peak.
16. **Time to peak deceleration as a percentage of total movement time (TPD%):** The amount of time taken from the start of the movement and the point where the deceleration peak occurs, expressed as percentage of the total movement time.
17. **Time to peak deceleration (TPD):** The amount of time taken from the start of the movement and the point where the deceleration peak occurs.
18. **Mean velocity (MV):** The mean velocity from the start of the movement and the end of the movement.

19. **Mean angular velocity (MAV):** The mean angular velocity from the start of the movement and the end of the movement.
20. **Mean acceleration (MA):** The mean acceleration from the start of the movement and the end of the movement.
21. **Mean angular acceleration (MAA):** The mean angular acceleration from the start of the movement and the end of the movement.
22. **Time lift (TL):** The amount of time from the beginning to the end of the lift phase (see Figure 28).
23. **Peak lift (PL):** The maximum value of the distance during the lift phase.
24. **Total movement (TM):** The sum of the movement time and the absolute reaction time (not expressed in percentage)
25. **Jerk score (JS):** A measure to quantify smoothness of the wrist path (Alberts et al. 2000, Teulings et al. 1997). Jerk score has to be normalized in time and amplitude to avoid the terrible increasing associated to movement duration (Schneider and Zernicke 1989). Then jerk score was computed using the following formula:

$$JS = \frac{1}{2} \times \int_{start\_mov}^{end\_mov} j^2(t) dt \times d^5 / l^2 \quad (5.10)$$

Where  $j$  is the jerk computed as described before (derivative of the acceleration),  $d$  is the movement duration (time movement) and  $l$  is the movement amplitude (distance peak between start movement and end movement). The definite integral from the start of the movement to the end of the movement is used to define a measure to quantify the smoothness of the wrist path during the movement.

The features described above are then normalised through being expressed as percentage. Those features not expressed as percentage (e.g. peak velocity) are transformed considering the maximum value of the feature contained in the data. For example the percentage of the peak velocity is computed in the following way:

$$Peak\ vel\ perc = \frac{peak\ vel}{max\ peak\ vel} * 100 \quad (5.11)$$

As explained before these 25 features are computed considering two different studies (Caselli et al. 1999, Alberts et al. 2000). In particular almost all the features relative to the distance, velocity and acceleration were computed in the Caselli's study (Caselli et al. 1999). The time lift and the peak lift were added. The features relative to the linear velocity and acceleration were computed also for the angular ones, using the findings of Alberts' study (Alberts et al. 2000) where some

differences between angular quantities are evidenced. Also the jerk score was replicated using the findings of Alberts' study.

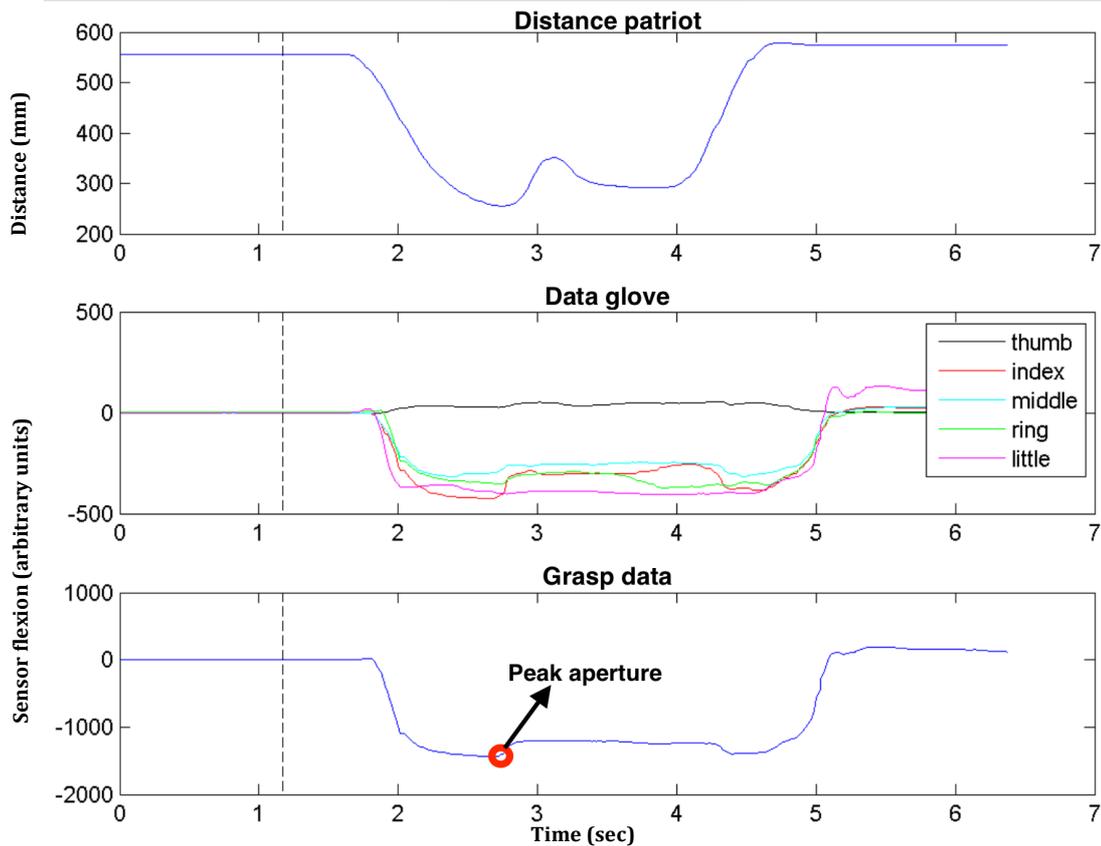
### **5.2.6 Grasping data**

The extraction of proper features from the data glove resulted to be very complicated because the data glove software provided flexion and extension data as a series of arbitrary integer values, without any measure unit. The value of each sensor increased when the finger flexed and decreased when it extended (often becoming a negative number).

The movements recorded by the sensors for each finger in some case were minimal and difficult to distinguish. For this reason, in order to have more readable data, the information from the middle, ring and little fingers were added to thumb and index finger data having an average of the fingers.

Peak aperture was estimated as the point in which the average of the fingers reached the minimum (fully extended). Then the time to peak aperture (TAP) was computed as the time between the start of the movement and the peak aperture point. TAP was the only feature that was possible to estimate from the data glove. It was not possible to calculate the amplitude of peak aperture, usually computed as index finger to thumb aperture, because the data glove did not give any clue about the distance between fingers. For the same reason it was not possible to establish when the grasp started, usually computed as the first point where the index starts to extend or better as the point where the distance between index and thumb starts to increase (Scarpa and Castiello 1994).

Figure 30 reports the grasp data of a trial made by a healthy subject, evidencing the time to peak aperture. This feature was not considered in the set of features computed because it is computed considering an average of the 5 sensors so we are not very confident about the precision of its value. Therefore the use of TAP in classification did not improve the results.



**Figure 30. Grasp data relative to a trial performed by a healthy subject. In the figure are reported the profiles of: distance patriot (upper), data glove reporting the values of each sensor (middle) and average of fingers computed as sum of all the sensor values (below). Peak aperture is evidenced.**

### 5.2.7 Description of the data

The data used are collected in the experiment described in section 5.2.3. Every subject did the experiment five times with the dominant hand and five times with the non-dominant hand in the four conditions described in section 5.2.3. In summary we have for each subject 10 reaching data, collected with the wrist sensor and 10 grasping data, collected with the glove sensors in each of the four different condition, for a total of 40 reaching data and 40 grasping data. The features described in section 5.2.5 are extracted for each reaching data collected in the four different conditions (40 for each subject). The 25 features extracted from each reaching data are treated as a single sample also if the reaching data belong to the same subject. The different repetitions of each subject (5 with dominant hand and five with non-dominant hand) are treated as single sample to have enough data to train the classifier. Dominant and non-dominant hand are treated in the same way putting them in the big data set. In summary for each subject there are 10 data reaching samples for each of the four different conditions, the features described in section 5.2.3 are extracted for each of the different samples and are inserted in the big data set in order to train and test the classifiers evolved.

## 5.3 RESULTS

### 5.3.1 Cases analysis

Before applying the evolutionary algorithms to the data, considering as inputs the 25 features described in section 5.2.5 it was decided to do a preliminary analysis. As explained in section 5.2.3 the experiment was undertaken in four different conditions: NAT (natural speed following a sound clue), VIS (natural speed following a light clue), MAX (as fast as they can following a sound clue), MEM (with the eyes closed remembering the position of the object, following a sound clue). Two different well known algorithm Support Vector Machine (SVM) (Durgesh and Lekha 2010) and Artificial Neural Network (ANN) (Maind and Wankar 2014) were used to find the best classifiers in each condition considered, considering as input the 25 features.. This analysis was done to understand if there is one case in which the difference among all classes considered is greater than the others. In other words we looked for the most discriminative visual condition. We decided to perform this analysis before applying CGP in order to limit the number of CGP simulations, which are very time consuming. Also the study of the classification results for each different condition could be very helpful to understand what condition is the best to appreciate the differences among the different classes considered. In table 2 the number of samples for each class in each case considered is reported. As explained is section 5.2.1, fifty-eight PD patients and 29 healthy subjects (Con) were tested. The patients are divided in 22 PD-NC, 23 PD-MCI and 10 PDD with three patients excluded in the categories but included in the overall patients. Each subject did the experiment 5 times for each hand in each visual condition considered (40 samples for subjects in total). The numbers reported in table 2 are fewer than expected because some data were discarded due to measurement errors or other protocol violations.

<b>Cases</b>	<b>Controls</b>	<b>PD</b>	<b>PD-NC</b>	<b>PD-MCI</b>	<b>PDD</b>
NAT	246	489	168	197	94
VIS	243	496	173	199	94
MAX	244	488	172	191	95
MEM	245	483	172	189	92

**Table 2. Number of samples considered for each class in each condition**

Given the classes in table 2, seven pairwise classification cases are considered:

- PD vs Con
- PD-NC vs Con
- PD-MCI vs Con
- PDD vs Con
- PD-NC vs PD-MCI

- PD-NC vs PDD
- PD-MCI vs PDD

Each dataset, comprising the two classes considered in each classification case, was divided into five folds and k-fold cross validation (k=5) was performed to generalize the results. In 5-fold cross validation, every time a different set of data for training, validation and test is considered, giving a good estimation of how the classifier generalise the results. Ten runs of the experiment were completed for statistical significance.

The SVM is configured with a linear kernel function and the solver used is SMO (Sequential minimal Optimization) (Fan, Chen and Lin 2005). These are the default settings in MATLAB. Different experiments were done using other kernel functions (Gaussian, polynomial) and a different solver (Iterative Single Data Algorithm) without an improving of the results, so we decided to use the default settings. The Neural network used was a feed-forward back propagation network with 20 hidden layers and a tan-Sigmoid transfer function. The number of hidden layers was chosen using an iterative method; we increased the number of hidden layers gradually registering the maximum improvement in the results when 20 is considered. A greater number of hidden layers did not improve the results so the hidden layer number was chosen equal to 20.

We choose as measure to evaluate the classifier obtained the Area Under ROC Curve AUROC (Hanley and McNeil 1982). The AUROC, as Hanley and McNeil demonstrated in their paper, represents the “probability of correctly ranking a (normal, abnormal) pair”. The ROC curve is a plot of the true positive rate on y-axis (TPR = true positive/number of positives) against the false positive rate on x-axis ( FPR = false positive/number of negatives ) considering different thresholds. The AUROC is then the area under the ROC curve. This measure is widely used for medical applications (van Erkel and Pattynama 1998, Lacy et al. 2013) and also to evaluate in general a machine learning algorithm (Bradley 1997).

In table 3 the classification results for SVM and ANN are reported considering all conditions together. While in tables 4, 5, 6 and 7 the results for each single condition are reported.

In each table the result reported for each case is in the form: mean  $\pm$  standard deviation, where the mean and the standard deviation of the AUROCs are computed across the ten runs of the experiment. Both values of train and test set are reported.

**ALL CONDITIONS (NAT, VIS , MAX, MEM).**

CASES	SVM		ANN	
	Test	Train	Test	Train
PD vs Con	$0.74 \pm 0.004$	$0.76 \pm 0.001$	$0.75 \pm 0.01$	$0.77 \pm 0.02$
PD-NC vs Con	$0.70 \pm 0.003$	$0.73 \pm 0.001$	$0.72 \pm 0.01$	$0.76 \pm 0.01$
PD-MCI vs Con	$0.77 \pm 0.004$	$0.79 \pm 0.001$	$0.79 \pm 0.02$	$0.82 \pm 0.01$
PDD vs Con	$0.86 \pm 0.004$	$0.88 \pm 0.001$	$0.88 \pm 0.01$	$0.90 \pm 0.01$
PD-NC vs PD-MCI	$0.71 \pm 0.004$	$0.75 \pm 0.001$	$0.70 \pm 0.01$	$0.73 \pm 0.02$
PD-NC vs PDD	$0.78 \pm 0.006$	$0.82 \pm 0.002$	$0.79 \pm 0.01$	$0.83 \pm 0.01$
PD-MCI vs PDD	$0.79 \pm 0.005$	$0.83 \pm 0.002$	$0.81 \pm 0.01$	$0.85 \pm 0.05$

**Table 3. SVM and ANN classification results for all conditions. AUROC mean and standard deviation across the ten runs is reported.**

**NAT (Natural speed).**

CASES	SVM		ANN	
	Test	Train	Test	Train
PD vs Con	$0.68 \pm 0.04$	$0.70 \pm 0.04$	$0.68 \pm 0.01$	$0.74 \pm 0.02$
PD-NC vs Con	$0.66 \pm 0.03$	$0.74 \pm 0.02$	$0.65 \pm 0.02$	$0.74 \pm 0.03$
PD-MCI vs Con	$0.68 \pm 0.03$	$0.75 \pm 0.03$	$0.70 \pm 0.02$	$0.77 \pm 0.02$
PDD vs Con	$0.74 \pm 0.03$	$0.86 \pm 0.02$	$0.83 \pm 0.03$	$0.89 \pm 0.03$
PD-NC vs PD-MCI	$0.62 \pm 0.02$	$0.72 \pm 0.02$	$0.61 \pm 0.02$	$0.71 \pm 0.03$
PD-NC vs PDD	$0.69 \pm 0.03$	$0.82 \pm 0.04$	$0.76 \pm 0.03$	$0.85 \pm 0.03$
PD-MCI vs PDD	$0.72 \pm 0.02$	$0.84 \pm 0.03$	$0.75 \pm 0.04$	$0.86 \pm 0.04$

**Table 4. SVM and ANN classification results for NAT condition. AUROC mean and standard deviation across the ten runs are reported.**

**VIS (Light clue)**

CASES	SVM		ANN	
	Test	Train	Test	Train
PD vs Con	$0.65 \pm 0.03$	$0.69 \pm 0.03$	$0.76 \pm 0.03$	$0.81 \pm 0.02$
PD-NC vs Con	$0.68 \pm 0.02$	$0.76 \pm 0.01$	$0.72 \pm 0.03$	$0.81 \pm 0.02$
PD-MCI vs Con	$0.77 \pm 0.02$	$0.83 \pm 0.02$	$0.82 \pm 0.02$	$0.88 \pm 0.02$
PDD vs Con	$0.81 \pm 0.01$	$0.88 \pm 0.01$	$0.84 \pm 0.02$	$0.91 \pm 0.02$
PD-NC vs PD-MCI	$0.68 \pm 0.02$	$0.78 \pm 0.02$	$0.67 \pm 0.04$	$0.78 \pm 0.04$
PD-NC vs PDD	$0.68 \pm 0.03$	$0.82 \pm 0.02$	$0.77 \pm 0.02$	$0.87 \pm 0.03$
PD-MCI vs PDD	$0.77 \pm 0.02$	$0.89 \pm 0.01$	$0.76 \pm 0.01$	$0.86 \pm 0.03$

**Table 5. SVM and ANN classification results for VIS condition. AUROC mean and standard deviation across the ten runs are reported.**

### MAX (Maximum speed).

CASES	SVM		ANN	
	Test	Train	Test	Train
PD vs Con	0.70 ± 0.06	0.72 ± 0.07	0.79 ± 0.01	0.84 ± 0.01
PD-NC vs Con	0.64 ± 0.03	0.70 ± 0.04	0.77 ± 0.01	0.84 ± 0.01
PD-MCI vs Con	0.78 ± 0.02	0.84 ± 0.01	0.79 ± 0.01	0.86 ± 0.02
PDD vs Con	0.92 ± 0.01	0.97 ± 0.003	0.95 ± 0.01	0.98 ± 0.01
PD-NC vs PD-MCI	0.62 ± 0.03	0.70 ± 0.04	0.69 ± 0.04	0.79 ± 0.05
PD-NC vs PDD	0.77 ± 0.03	0.86 ± 0.04	0.84 ± 0.03	0.92 ± 0.03
PD-MCI vs PDD	0.74 ± 0.04	0.86 ± 0.04	0.84 ± 0.02	0.90 ± 0.02

**Table 6. SVM and ANN classification results for MAX condition. AUROC mean and standard deviation across the ten runs are reported.**

### MEM (Eyes closed).

CASES	SVM		ANN	
	Test	Train	Test	Train
PD vs Con	0.81 ± 0.01	0.86 ± 0.002	0.80 ± 0.01	0.84 ± 0.02
PD-NC vs Con	0.78 ± 0.01	0.86 ± 0.003	0.77 ± 0.02	0.84 ± 0.03
PD-MCI vs Con	0.81 ± 0.01	0.87 ± 0.002	0.79 ± 0.02	0.86 ± 0.02
PDD vs Con	0.94 ± 0.01	0.99 ± 0.002	0.94 ± 0.01	0.98 ± 0.01
PD-NC vs PD-MCI	0.78 ± 0.01	0.83 ± 0.001	0.72 ± 0.04	0.82 ± 0.03
PD-NC vs PDD	0.82 ± 0.01	0.89 ± 0.002	0.80 ± 0.03	0.89 ± 0.03
PD-MCI vs PDD	0.80 ± 0.01	0.89 ± 0.002	0.81 ± 0.03	0.90 ± 0.03

**Table 7. SVM and ANN classification results for MEM condition. AUROC mean and standard deviation across the ten runs are reported.**

Examining all the results in the tables 3-7, MEM seems to be the best discriminating condition. This result is expected, because it is reasonable to think, that with the eyes closed the difficulties increases for the patients and also cognitive impairments should be more noticeable.

There are some interesting considerations examining all the results. We can notice that we have the worse results for the conditions NAT and VIS. These conditions do not seem to evidence the differences in performing the experiment neither between patients and controls, nor between the cognitive subgroups of patients. A possible explanation of these findings could be that at natural speed the patients are relaxed and performed the experiment in a similar way as the controls. For the VIS condition surprisingly, the finding that partial visual feedback affects badly the patients performance in reach and grasp (Schettino et al. 2003) is not confirmed. An explanation of this could be that maybe the room was not dark enough to remove the visual feedback affecting patient's performance.

The results of MAX condition are very interesting. ANN can evolve classifiers with performances comparable to MEM condition. In “PD-NC vs PDD”, “PD-MCI vs PDD” and “PDD vs Con” the classifiers evolved using ANN have better performance on the test set than those evolved for the MEM condition. In MAX condition, ANN outperforms SVM. Then from this results ANN seems to be the best technique to evolve classifiers especially for MAX condition.

MEM condition can be considered the best condition for two main reasons:

1. Both SVM and ANN evolved classifiers with comparable performances, so we can be more confident about the results;
2. The classifiers evolved have the best overall performances compared to the other conditions (excepted for the cases described above).

In this study we consider only the data of MEM condition, evaluating it as the best condition to distinguish among the class considered.

### **5.3.2 CGP classification results**

In this section the classification results of CGP are reported. Only the MEM condition is considered for the reasons explained in the previous subsection.

The geometry of the programs in the population of CGP (chromosomes) has seventy-five internal nodes (three times the number of inputs) with a function set of eleven mathematical functions (+, -, \*, /, mean, min, max, division rest, sin, cos, tan), 25 inputs (the features described in section 5.2.5) and one output. The number of internal nodes equal to three times the number of inputs was chosen through experimentations using an iterative method, a greater number of internal nodes incremented the complexity but did not improve the results. As function set we considered a basic set of mathematical functions since we did not have specific knowledge of the problem to determine specific functions. Each output is considered a positive response if the value is greater than a certain threshold, negative otherwise; 18 thresholds are considered here, in multiples of 5 (5, 10, 15...., 90). We chose to use 18 thresholds to facilitate the calculation of AUROC. The values of the threshold are chosen taking account of the input values expressed in percentage and then comprised between 0 and 100. For each threshold the true positive rate and false positive rate is computed, then the ROC curve is depicted using these two computations for each threshold. The positive class is the first in the classification cases; for example, for the case “PD-NC vs PD-MCI”, the positive class is PD-NC. At each generation the fittest chromosome is selected and the next generation is formed with its mutated versions (mutation rate=0.07). Evolution is stopped when 50000 iterations are reached or when the over fitting occurs.

In table 8 all the results are summarized. The SVM and ANN results are the same reported in table 7, they are used as comparisons in order to evaluate CGP results. ANN and SVM are used as baseline methods because are two very well-known techniques and were applied before to medical diagnosis. ANN was used before to diagnose Parkinson's disease and other kinds of dementia (Geman 2013) and also for the diagnosis of different diseases such as colorectal cancer (Spelt et al. 2012), multiple sclerosis lesions (Mortazavi, Kouzani and Soltanian-Zadeh 2012), colon cancer (Ahmed 2005) and pancreatic disease (Bartosch - Härlid et al. 2008). SVM was used to identify potential biomarker in psychiatric and neurological disorders (Orru et al. 2012) but also for the diagnosis of different diseases such as cancer (Wang and Huang 2011), heart disease (Ghumbre, Patil and Ghatol) and diabetes (Zhang et al. 2017)

The classification cases considered here are the same reported in subsection 5.3.1 and the 5-folds cross validation was performed. As in tables 3-7, the results are reported in the form of mean  $\pm$  standard deviation, where the mean and the standard deviation of the AUROCs are computed across the ten runs of the experiment. Both results for train and test set are present for each technique.

CASES	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
PD vs Con	0.81 $\pm 0.01$	0.86 $\pm 0.002$	0.80 $\pm 0.01$	0.84 $\pm 0.02$	0.79 $\pm 0.01$	0.82 $\pm 0.01$
PD-NC vs Con	0.78 $\pm 0.01$	0.86 $\pm 0.003$	0.77 $\pm 0.02$	0.84 $\pm 0.03$	0.76 $\pm 0.01$	0.81 $\pm 0.02$
PD-MCI vs Con	0.81 $\pm 0.01$	0.87 $\pm 0.002$	0.79 $\pm 0.02$	0.86 $\pm 0.02$	0.78 $\pm 0.01$	0.81 $\pm 0.03$
PDD vs Con	0.94 $\pm 0.01$	0.99 $\pm 0.002$	0.94 $\pm 0.01$	0.98 $\pm 0.01$	0.93 $\pm 0.01$	0.95 $\pm 0.01$
PD-NC vs PD- MCI	0.78 $\pm 0.01$	0.83 $\pm 0.001$	0.72 $\pm 0.04$	0.82 $\pm 0.03$	0.65 $\pm 0.02$	0.73 $\pm 0.02$
PD-NC vs PDD	0.82 $\pm 0.01$	0.89 $\pm 0.002$	0.80 $\pm 0.03$	0.89 $\pm 0.03$	0.78 $\pm 0.01$	0.82 $\pm 0.03$
PD-MCI vs PDD	0.80 $\pm 0.01$	0.89 $\pm 0.002$	0.81 $\pm 0.03$	0.90 $\pm 0.03$	0.78 $\pm 0.02$	0.84 $\pm 0.02$

**Table 8. Classification results for MEM condition. AUROC mean and standard deviation across the ten runs are reported.**

As the test set results are most useful in evaluating the potential of the classifier to generalise the unseen data, from this point forward discussion regarding the results will refer solely to the test set. Examining the results in table 8, unfortunately the classifiers evolved by CGP seem to be the worse for each case. However CGP does give an advantage to us to understand the most important inputs for each case,

because the programs do not necessarily use all the inputs considered. In addition the programs evolved by CGP are simple mathematical expressions easy to implement (details in section 5.3.3).

In all the classification cases containing controls (“PD vs Con”, “PD-NC vs Con”, “PD-MCI vs Con” and “PDD vs Con”) the classifiers evolved by CGP are comparable to the classifiers evolved by the other techniques. In fact, in all these cases, the results of CGP differ a maximum for 3% with respect to the best method (SVM or ANN). We can conclude that in all cases comprising the healthy subjects, CGP is able to evolve classifiers with approximately the same performance of those evolved using SVM and ANN.

Unfortunately the situation changes when the comparisons among the subgroups of patients are considered. In these cases CGP evolved classifiers with worse performances than SVM and ANN.

In the case “PD-NC vs PD-MCI” SVM evolved the best classifiers reaching 0.78, followed by ANN reaching 0.72 and both methods outperforms CGP which evolves classifiers reaching 0.65. In this case CGP does not seem the best approach to evolve classifiers; it is preferable to use SVM that has very good results.

Considering “PD-NC vs PDD” and “PD-MCI vs PDD” CGP results differ respectively of approximately 4% and 3% with respect to the best method. In these cases the classifiers evolved by the CGP was more comparable to those evolved using the other methods.

We can conclude that, if CGP is not the best method, in many cases it evolves classifiers comparable to the other methods and has two possible advantages: selecting only the most useful inputs for the problems and evolving programs that are simple mathematical expressions and therefore are easy to implement. In the next subsection we will give details about the mathematical expression derived by the evolved classifier and the network diagram associated

### **5.3.3 CGP classifier Mathematical expressions and Network Diagrams**

As shown in section 2.4 one of the main advantages of CGP is that the evolved classifier can be represented as a group of standard mathematical expressions and corresponding directed graphs (figure 14). In figure 14 it is clear that each output can be represented as a simple mathematical expression using the directed graph representing the classifier. In our case the classifier evolved has only one output and can be represented as a single mathematical expression. In figure 31 a network diagram corresponding to the best classifier evolved for the case “PDD vs Con” is reported. Cross-validation is used, so the best classifier considered is the best across all the folds and the generations. The classifier considered has an AUROC equal to 0.98 for the test set. In figure 31 only the active nodes are reported in order to make the diagram clearer. We can note that only ten of the twenty-five inputs contribute to the output computation. The order of the input is reported in section 5.2.5

(starting from 0 instead of 1); following this order the inputs used by the classifier in this case are: movement time (MT), reaction time as percentage of total movement time (RT%), time to peak acceleration as percentage of total movement time (TPA%), time to peak angular acceleration (TPAA), peak deceleration (PED), time to peak deceleration as percentage of total movement time (TPD%), mean velocity (MV), mean angular velocity (MAV), mean acceleration (MA) and time lift (TL). The study of the inputs used can help to understand what the most useful features for each classification case are. Unfortunately, identifying the most useful features for each classification case was not possible in this context. In fact there was not a common set of inputs used by the best classifier for each case across the folds of cross-validation. For this motivation it is not possible to make any conclusion about the most important features for each classification case. However, the diagram reported in figure 31 helps to understand what happens inside the classifier. The representation of the classifier is easy to understand and gives a better insight about the data used. CGP was used principally for this motivation, the representation of the classifiers evolved permits a better understanding of how the inputs are used and the output can be translated in a single mathematical expression, facilitating the implementation. Figure 31 can be represented easily with a mathematical expression following the oriented edges of the graph (figure 14). The mathematical expression corresponding to the diagram is expression 5.12. In the expression the input numbers are replaced with the short name of the inputs (section 5.2.5) to clarify the meaning of the equation. The expression is complex but helps to understand how the equation can be easily derived from the diagram.

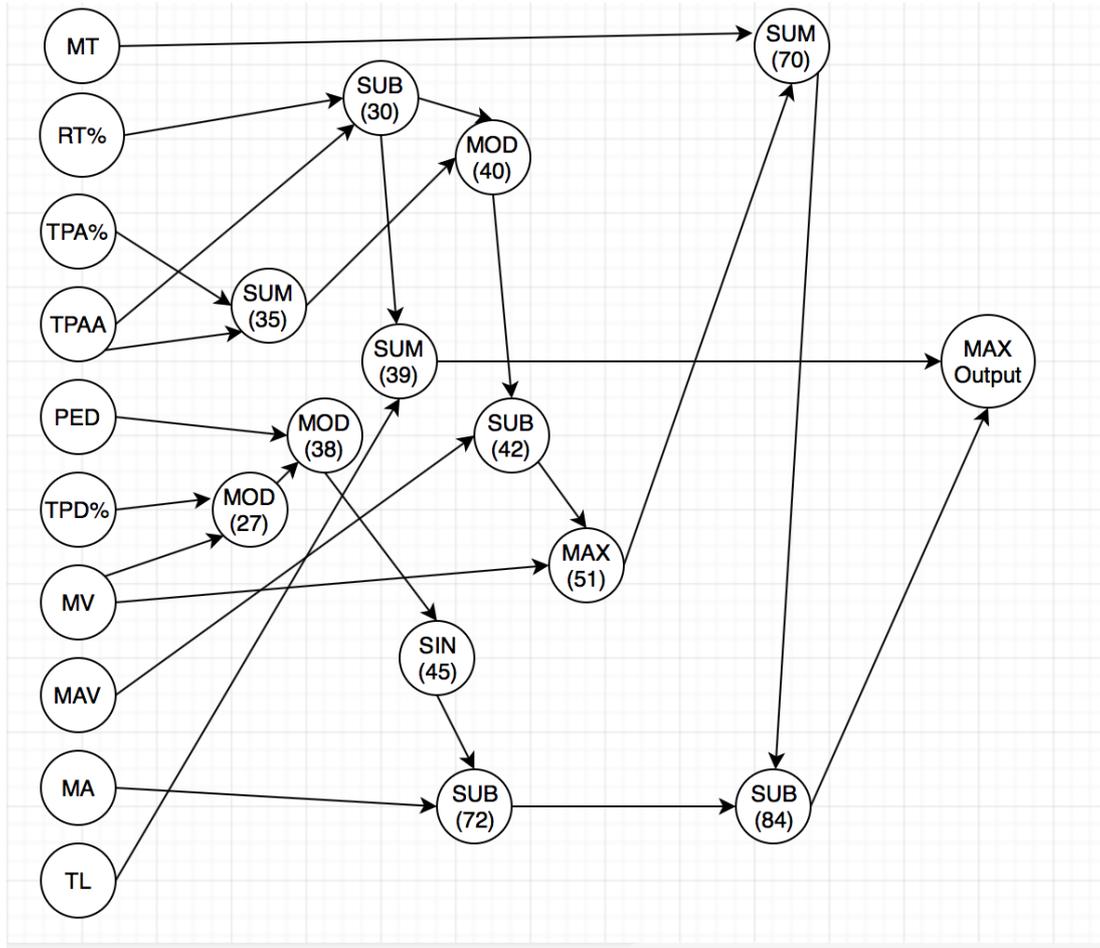


Figure 31. Network diagram of the best classifier evolved for the case PDD vs Con. Only the active nodes are reported.

$$out = \max \{ [(MT + \max ((MAV - (\text{mod}((RT\% - TPAA), (TPA\% + TPAA))), MV) - (MA - \sin(\text{mod}(\text{mod}(TPD\%, MV), PED))), [(RT\% - TPAA) + TL]] \}. \quad (5.12)$$

## 5.4 CONCLUSIONS AND FUTURE WORKS

Examining table 8 interesting conclusions on each single case can be made. Let's start with the different comparisons between patients and controls. We can see that the classification between PD (all PD patients together) and Controls is around 0.79-0.81. The comparisons between subgroups of patients and controls, the situation is different. For "PD-NC vs Con" the mean AUROC of the generated classifiers is 0.76-0.78, for "PD-MCI vs Con" it is 0.81-0.78 and for "PDD vs Con" it is 0.93-0.94. As expected, the classification between PD-NC and Controls is less certain than the other subgroups cases. The best classification is for "PDD vs Controls" in which the generated classifiers have a very high AUROC. Again, this result is expected: based on conventional clinical assessment, it is easier to distinguish between PDD and Controls than between PD-NC and Controls, or PD-MCI and Controls. Therefore, it can be proposed that the PD-NC class lowers the

overall classification accuracy between PD and Controls, whilst the PDD class raises the classification accuracy.

Comparisons of classification results for PD patient subgroups are interesting. The most challenging classification is between PD-NC and PD-MCI. This is expected because we know that patients from the two classes can perform the movement in a similar way and can be difficult to distinguish clinically. In this case CGP generated classifiers with a mean AUROC of 0.65 and it is outperformed by both ANN and SVM. SVM generates the best classifiers with a mean AUROC of 0.78. The differentiation between PD-NC and PDD, and between PD-MCI and PDD is more pronounced. This result was also expected because we know that PDD patients have greater difficulties with respect to the other subgroups in performing the movement. In fact, we note that for “PD-NC vs PDD” the mean AUROC of the generated classifiers is 0.78-0.82 while for the case “PD-MCI versus PDD” it is 0.78-0.81.

For all the comparisons between patients and controls CGP evolved classifiers comparable to the other approaches while for the comparisons among the patients the differences among the classifiers evolved by CGP and the other methods, in some cases, are more evident. In particular in the case “PD-NC vs PD-MCI” the classifiers evolved by CGP are outperformed by the other methods.

In the study considered there are two limitations that could influence the results. Firstly, the patients examined were on medication to limit motor symptoms of PD, which may affect their performance. The second limitation of this study is the relatively small number of the patients measured and the consequent way in which the different repetitions were considered. Each subject performs the experiment five times with each hand. So for each subject there are ten repetitions that are considered in the dataset as ten different samples. The data available resulted less than those recorded for measurement errors and protocol violations (table 2). If we only considered the mean of the repetitions, the data available would be insufficient to evolve a classifier. The strategy to consider repetitions as different samples could influence the results because the same subject can perform the experiment in a similar way, reducing the variability. However, patients showed a high variability in performing the experiment, so including repetitions was justified. In future, a new memory guided reach experiment including more subjects would be useful to overcome this limitation.

Another point of discussion is the use of the grasping data. They are not used for the reasons explained in section 5.2.6, but could be useful to find new features to improve the classification. In the future it would be useful finding a way to use these data, it is suggested to consider a new reach and grasp experiment using two Polhemus Patriot sensors (described in section 5.2.2): one placed on the index finger and another placed on thumb finger, in order to measure the distance between the fingers.

Moreover the results using CGP could be improved when considering other features or using the raw data values as inputs. In particular, the use of the raw data has a big potential, because the whole signal is considered instead of a summary

represented by the features. In conclusion, the results of this study are promising and can be used as a starting point in the classification of PD patients and controls but also in the detection of different cognitive impairments considering the “reach and grasp” task.

In the next chapter the wavelet transform are introduced. We will use this technique to pre-process the hand-opening data (details in Chapter 7).

## CHAPTER 6: WAVELET TRANSFORM ANALYSIS

In this chapter an introduction about wavelet transform analysis is given. This technique is used to pre-process the EMG of dystonia patients and healthy subject, recorded in the second part of the study (details in chapter 7). The wavelet transform analysis is a technique used to perform time-frequency analysis. The time-frequency analysis gives information about all the frequencies contained in the signal and the time at which they occur. The Fourier transform differs from the wavelet transform in the fact that it has only frequency resolution and no time resolution. In order to find the power spectra, the amplitude of all the existing sine functions at various frequencies are identified through the whole duration of the signal (Le Van Quyen and Bragin 2007). This means that, we know all the frequency components in the signal, but the time in which they occur is unknown (Haddad and Serdijn 2009). Fourier transform is defined as a frequency–amplitude decomposition (Poularikas 2010).

Time-frequency analysis is more suitable than just frequency analysis made with the Fourier transform when the signal examined is oscillating and is dynamic with frequency changes in time (Le Van Quyen and Bragin 2007). In these situations it is important not only knowing the frequency component but also decoding the oscillations separating them in elementary units well-defined in frequency and time (Le Van Quyen and Bragin 2007). Wavelet transform is used to perform such time-frequency analysis.

All the concept described in this chapter are taken by Addison's book (Addison 2002) unless otherwise specified.

In wavelet transform analysis the functions used are called wavelets because they have a waveform. The wave-like scalable function is localized in time and frequency. This kind of function is more suitable in representing abruptly signal changes in time with respect to sine functions used in Fourier transform (Le Van Quyen and Bragin 2007). There are two main operations with the wavelet: translation and scaling. With translation the wavelet is moved to a different location, with scaling it is stretched or squeezed. In figure 32a scaling and translation operations are represented.

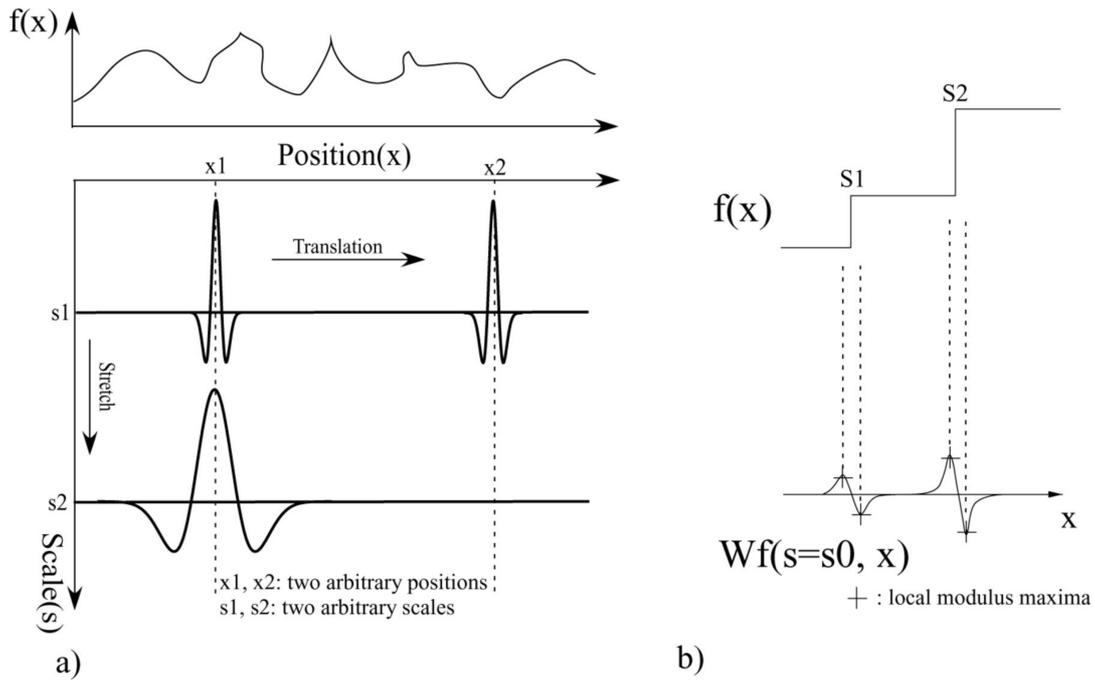
Using wavelets a time domain signal can be transformed into a two dimensions time-frequency domain signal, useful to evidence features. The transformation of the signal is called wavelet transform. In mathematical terms, wavelet transform is defined as the convolution of the wavelet function with the signal. From a practical point of view, we can say that the wavelet transform has a big value when wavelet matches the shape of the signal for a specific scale and location; instead it has a low value when this does not happen. In figure 32b the value of the wavelet transform for a certain scale  $s_0$  is reported. It is possible to note how for the smaller scales the wavelet converges to the function singularities S1 and S2.

The transform value is located in a two dimensional space, called transform plane, using location and frequency (derived from scale). The representation of the transform plane is illustrated in figure 33.

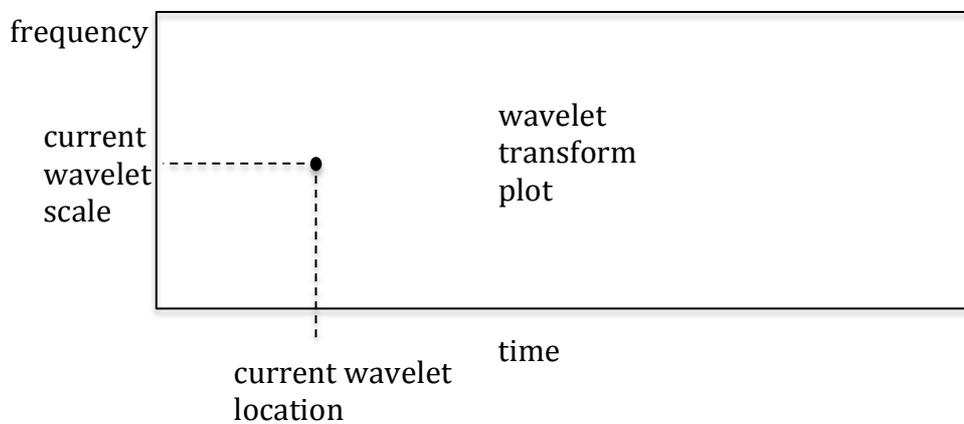
The wavelet transform is computed at various locations and scales, filling up the transform plane. The wavelet transform gives information about time and frequency at the same time, giving all the frequencies contained in the signal along with the time when they occur. Wavelet transform is in practice computed for each scale of interest across all time points. This is done decomposing the signal of interest into several parts and then analysing the parts separately (Haddad and Serdijn 2009). The several parts are represented by the different locations and scale contained in the transform plane: the location is related to the time while the scale is inversely proportional to the frequency. Dilation in time domain corresponds to lengthening time periods and a lowering of the associated frequency. The scaling operation is used to examine the signal at various levels of details: wavelet contraction is associated to the increasing depth of focus, passing from coarser (i.e. low- frequency) to finer (i.e. high-frequency) signal structures (Le Van Quyen and Bragin 2007). The way in which the wavelet transform can decompose the signal in the different parts will be further explained in the next sections.

There are two kinds of wavelet transform: the continuous wavelet transform (CWT) and the discrete wavelet transform (DWT). The CWT is used when the transform plane is filled in a continuous or smoothed way, while the DWT is used when it is filled in discrete steps. In the following subsections both kinds of wavelet are discussed but CWT is discussed in more detail because this is the approach chosen for the work described in this thesis.

In this work we decided to use wavelet transform instead of Fourier transform because it has been shown to be a very promising mathematical tool (Rioul and Vetterli 1991, Daubechies 1992, Mallat 1999), particularly for local analysis of non-stationary and fast transient signals (as medical signals), due to its good estimation of time and frequency localizations (Haddad and Serdijn 2009).



**Figure 32. Scaling and translation of a wavelet plus wavelet transform for a certain scale (a) Scaling and translation of a wavelet (here the second derivative of the Gaussian) along a signal (function)  $f(x)$ . (b) Up: a function  $f(x)$ , down: amplitude of the wavelet transform of function  $f(x)$ , along the  $x$ -axis, at a certain scale,  $s_0$ . The original wavelet is the second derivative of the Gaussian. Source: (Enescu, Ito and Struzik 2006)**



**Figure 33: Transform plot of the wavelet transform. Source: (Addison 2002)**

This chapter is divided in three main sections.

In section 6.1 the continuous wavelet transform (CWT) is described illustrating also some applications. Section 6.1 is divided in three subsection: subsection 6.1.1 contains the description of the two wavelets spectrum, the Fourier spectrum and the wavelet-based energy spectrum, in subsection 6.1.2 the Heisenberg boxes of the wavelets transform are discussed and in subsection 6.1.3 two kinds of wavelets transform are discussed: Morlet wavelet and generalized Morse wavelets (the approach used in this work).

In section 6.2 a brief introduction of discrete wavelet (DWT) transform is given. Subsection 6.3 contains a summary highlighting the key-points described in the chapter.

## 6.1 THE CONTINUOUS WAVELET TRANSFORM

A wavelet is a function that has to satisfy certain mathematical constraints. There are many different wavelets; the choice depends on the signal to transform and on the application. The wavelet chosen is called the mother wavelet and it can be altered with two main operations: translation (movement along time axis) and dilation (spreading out).

The wavelet function  $\psi(t)$  has to satisfy three mathematical criteria:

1. finite energy:

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \quad (6.1)$$

where E is the energy of the function computed as the integral of its squared magnitude, the vertical brackets represent the modulus operator giving the magnitude of  $\psi(t)$ . If  $\psi(t)$  is a complex function both real and complex part have to be used to compute its magnitude

2.  $\psi(t)$  must have zero mean, equivalent to no zero frequency component,  $\hat{\psi}(0) = 0$  where  $\hat{\psi}(f)$  is the Fourier transform of  $\psi(t)$ :

$$\hat{\psi}(f) = \int_{-\infty}^{\infty} \psi(t) e^{-i(2\pi f)t} dt \quad (6.2)$$

Then the following condition has to be satisfied:

$$C_g = \int_0^{\infty} \frac{|\hat{\psi}(f)|^2}{f} df < \infty \quad (6.3)$$

where  $C_g$  is called admissibility constant and has a value depending on the chosen type of wavelet.

3. In case the wavelet is a complex function (e.g. Morlet wavelet) the Fourier transform has to be real and dissolving for negative frequencies.

In figure 32 the operations of dilatation and translation are illustrated. The mother wavelet used in this figure is called Mexican hat and is the second derivative of a Gaussian distribution function  $e^{-t^2/2}$ :

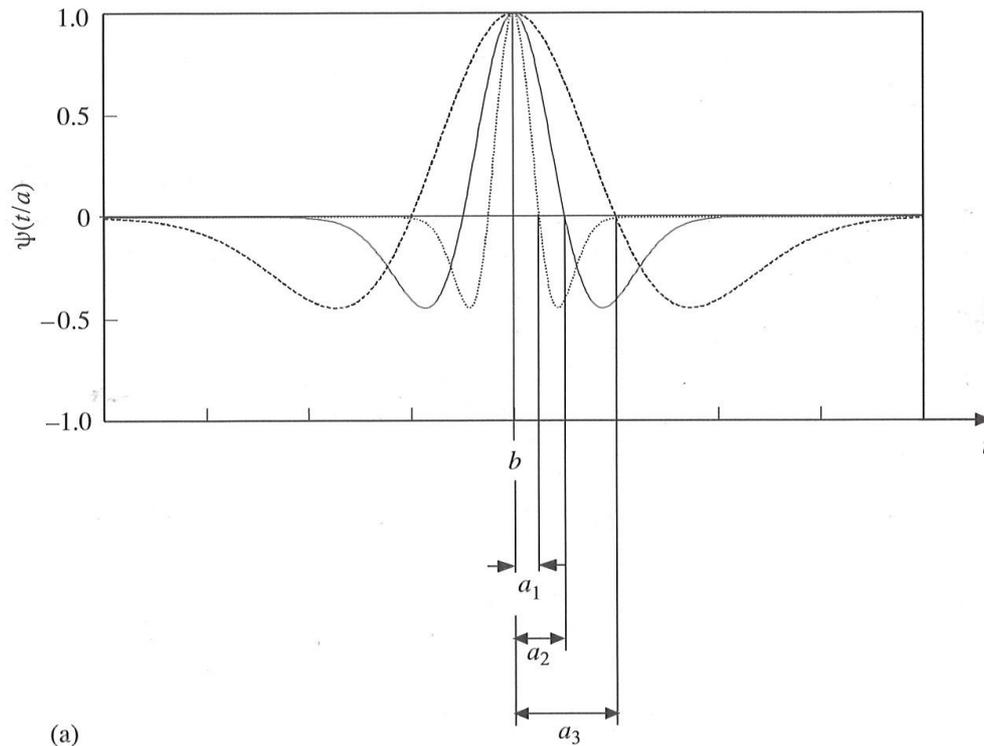
$$\psi(t) = (1 - t^2)e^{-t^2/2}. \quad (6.4)$$

The dilatation parameter  $a$  is used to control the dilation and the contraction of the wavelet, while the translational parameter  $b$  is used for the wavelet movements along the time axis (figure 34). Both parameters  $a$  and  $b$  can be added to the wavelet definition representing all the shifted and dilated version of the mother wavelet as  $\psi[(t - b)/a]$ . The expression of the mother wavelet  $\psi(t)$  is found when  $b = 0$  and  $a = 1$ .

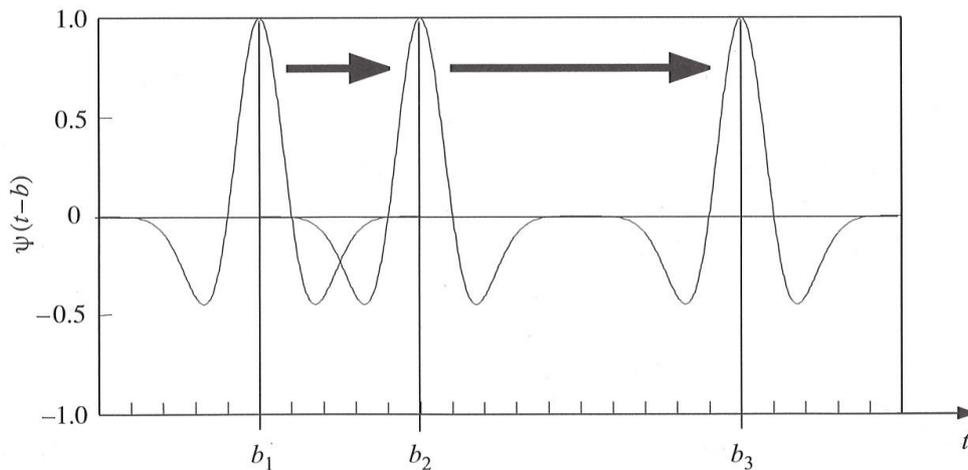
The wavelet transform of a continuous signal  $x(t)$  with respect to the wavelet function is defined as:

$$T(a, b) = w(a) \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t - b}{a} \right) dt \quad (6.5)$$

Where  $w(a)$  is called the weighting function and its value is usually equal to  $1/\sqrt{a}$  for energy conservation (at each scale the wavelets have same energy). The complex conjugate of the wavelet function is indicated with asterisk. In equation 6.5 there is a product between the signal and the wavelet integrated over the signal range. This operation in mathematics is called convolution.



(a)



(b)

**Figure 34: Dilatation and translation of a wavelet.**

**a) Stretching and squeezing wavelet (dilation) using three different scales  $a_1, a_2$  and  $a_3$  defined as  $a_1 = a_2/2$ ;  $a_3 = a_2 \times 2$ , b) moving a wavelet: translation three different locations are considered  $b_1 < b_2 < b_3$ . Source: (Addison 2002)**

The convolution, described in equation 6.5, has a positive contribution when both the wavelet and the signal have the same sign, while it has a negative contribution when they have opposite sign.

In figure 35 an example illustrates this concept. In figure 35(a) a wavelet of a specific dilation and in a specific location is described. In regions A and B there is a positive contribution to the integral (wavelet and signal have same sign), while in

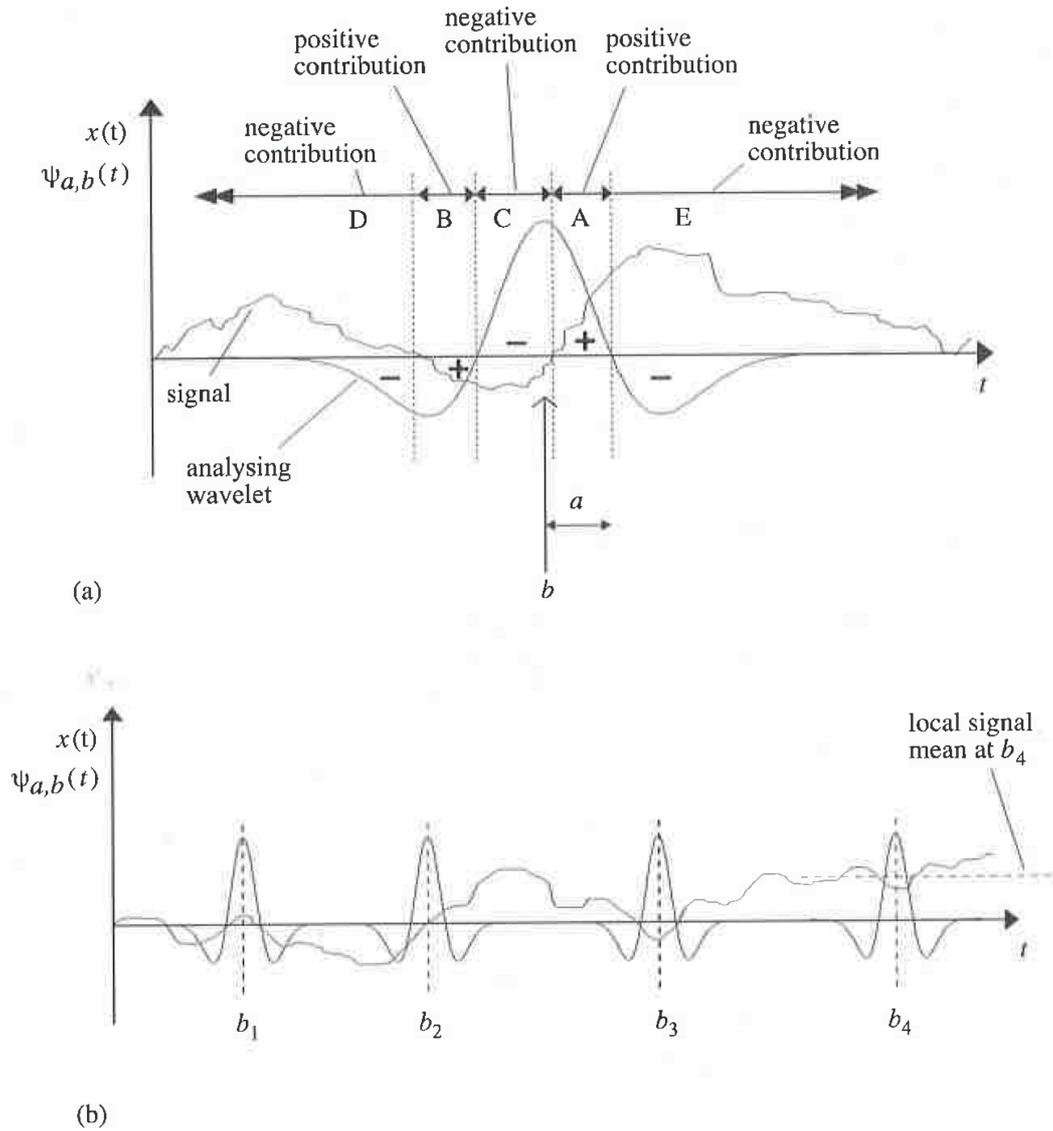
regions C, D and E there is a negative contribution to the integral (wavelet and signal have opposite sign).

In figure 35(b) there is a wavelet of a fixed dilation in four different locations. At the first location  $b_1$  the positive and the negative parts of the wavelet are almost coincident with those of the signal giving a large positive value of  $T(a, b)$ . At the second location  $b_2$  the positive and the negative contributions are almost the same deleting each other and resulting in a value near to zero for  $T(a, b)$ . At the third location  $b_3$ , signal and wavelet are out of phase resulting in large negative value for  $T(a, b)$ . At the fourth location  $b_4$  again the wavelet and the signal are out of phase, but this time the signal portion in proximity of the wavelet contains a local minimum component. This component contributes in the same way to positive and negative values of  $T(a, b)$ . Then only the local signal feature is highlighted and the mean is discarded, resulting in a negative value for  $T(a, b)$ . The process described is the way in which the wavelet identifies “coherent structures” in a time signal for different scale. It is repeated for all a scales moving the signal in each different  $b$  location, until all the wavelet coherent structures within the signal from the largest to the smallest for each scale are identified.

CWT is used for pattern recognition (Szu, Telfer and Garcia 1996, Qiao, Esmaily and Melhem 2012) because it has the ability to decompose complex signal and patterns into elementary forms.

There are different medical applications of CWT: for instance Iranmanesh and colleagues used CWT, and in particular the Mexican hat, to detect spikes due to epilepsy in electroencephalography (EEG) (Iranmanesh and Rodriguez-Villegas 2017) while Addison and colleagues used a Morlet wavelet transform to detect ventricular fibrillation in electrocardiogram (ECG) (Addison et al. 2002).

An important property of CWT the resistance to signal noise is highlighted in (Slavič, Simonovski and Boltežar 2003). Slavič and colleagues used Gabor wavelet transform to describe damping. Gabor wavelet (Carmona, Hwang and Torresani 1998) is a complex function characterized by a parameter  $\partial$ , with similar form of the Morlet wavelet described in section 6.1.3. Slavič and colleagues demonstrated that a large  $\partial$  parameter gives a slightly better resistance to noise. They demonstrated the advantages of using the amplitude and phase methods, both of which provide information about the instantaneous noise.



**Figure 35: Wavelet interrogation of the signal.**  
**a) The wavelet at specific dilatation and location on the signal. b) The wavelet of fixed dilatation at four different locations. Source: (Addison 2002)**

### 6.1.1 Wavelet energy

The wavelet has two different energy spectra: the Fourier spectrum and the wavelet-based spectrum (scalogram). These two spectra are defined in detail as followings:

1. The Fourier spectrum of a mother wavelet is defined as the plot of squared magnitude of the wavelet Fourier transform against wavelet frequency:

$$E_F(f) = |\hat{\Psi}(f)|^2 \quad (6.6)$$

where  $\widehat{\psi}(f)$  is the Fourier transform of the wavelet, the vertical brackets represent the module operator and  $E_F(f)$  is the Fourier spectrum of the mother wavelet  $\psi(t)$ . A wavelet is a bandpass filter, then only the signal components frequencies within a certain range of frequencies (passband) can go through it. These signal components are in proportions characterized by the energy spectrum of the wavelet defined in equation 6.6. The passband centre frequency of the mother wavelet,  $f_c$  is the standard deviation of the energy spectrum about the vertical axis and is defined as:

$$f_c = \sqrt{\frac{\int_0^\infty f^2 |\widehat{\psi}(f)|^2 df}{\int_0^\infty |\widehat{\psi}(f)|^2 df}} \quad (6.7)$$

The wavelet scale  $a$  is inversely proportional to  $f_c$  and to all mother wavelet characteristic frequencies. This is an obvious consideration because it is clear that a dilation in time domain means lengthening time periods corresponding to a lowering of associated frequencies. Wavelet Fourier spectrum is useful in visualising the wavelet band pass filter behaviour.

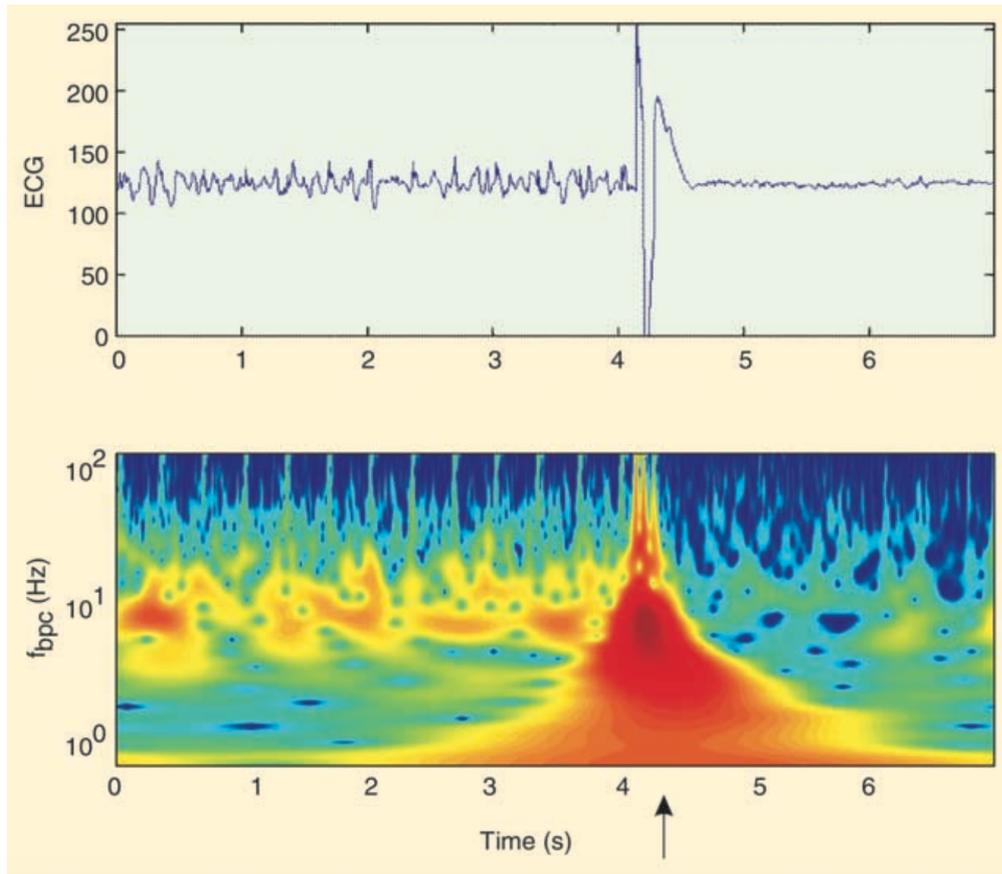
2. The wavelet-based energy spectrum is determined for a specific scale, location and signal  $x(t)$  as:

$$E(a, b) = |T(a, b)|^2 \quad (6.8)$$

where  $T(a, b)$  is the wavelet transform defined in equation 6.5. The plot of  $E(a, b)$  is called scalogram and all the function that differ from  $|T(a, b)|^2$  by a multiplicative constant are called scalogram. The scalogram is often expressed in a logarithmic  $a$  scale axis.

The scalogram is very useful in highlighting the location and the scale of dominant features in the signal. The following examples are useful to clarify this concept.

In figure 36 a human ECG containing a defibrillation shock event along with its scalogram is illustrated. Looking at the scalogram it is possible to notice the high frequency spike, evidenced with an arrow, preceding the traumatic event.



**Figure 36. Top: 7 seconds of human ECG containing a defibrillation shock event. Bottom: scalogram corresponding to ECG signal. Notice the high frequency spike in the scalogram before the shock event. Source: (Addison et al. 2002)**

In figure 37 the aorta pressure trace of a pig (top) showing an irregular activity of the ventricular muscle, which is obscured in the ECG (middle), is shown. At the bottom there is the wavelet scalogram obtained using a Morlet wavelet, explained in the subsection 6.1.3. The high frequency pulses in the scalogram (highlighted by white arrows) correspond approximately to the aortic pressure pulses. In this case then the wavelet scalogram is useful to identify the location of the aortic pressure pulses.

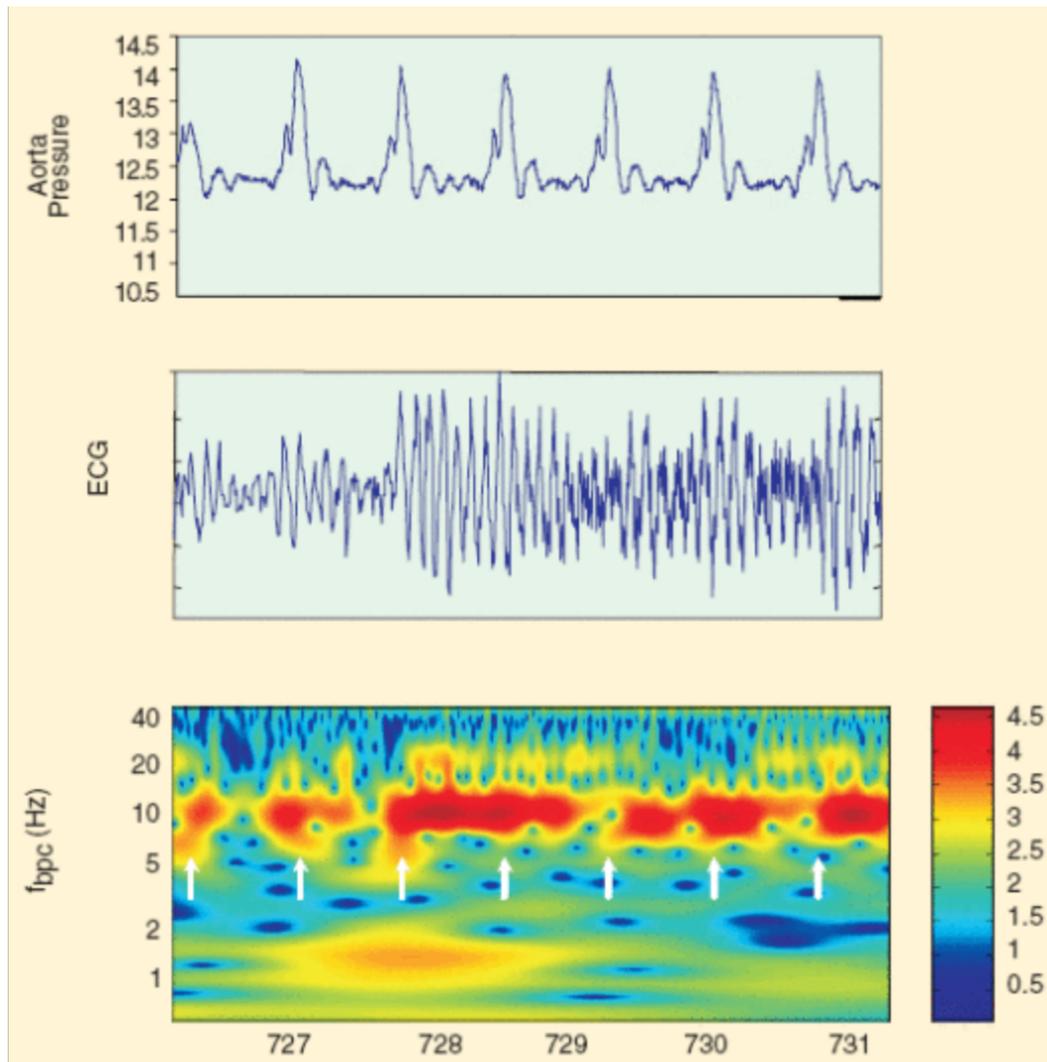


Figure 37. The aorta pressure trace (top) with the ECG (middle) and the corresponding wavelet energy plot (bottom) using the Morlet wavelet. The white arrows mark high-frequency “pulses” visible in the scalogram and corresponding approximately to the aortic pressure pulses. Source (Addison et al. 2002)

### 6.1.2 Heisenberg box wavelet transform

The wavelet transform, as explained so far, gives a representation in time and frequency using a two-dimensional transform. The Heisenberg uncertainty principle expresses the fact that when you are dealing with two dimensions, you cannot have the same resolution precision for both of them (Busch, Heinonen and Lahti 2007). In physics this can be translated as the impossibility to know at the same time where a particle is and how fast it is. The Heisenberg uncertainty principle then expresses a trade-off saying that if you want a more precise particle position you have to be less certain about its speed and vice versa.

This principle is applied also to the wavelet transform: a contraction in time of the wavelet corresponds to a wavelet containing higher frequencies with a wider spread and vice versa (Addison 2002). Heisenberg boxes are used to show the resolution

uncertainty in the two dimensions. Figure 38 shows some Heisenberg box examples illustrating the Heisenberg uncertainty principle.

In the middle of the figure three Morlet wavelet at three different scales are reported (only the real parts are shown), while in the bottom the energies densities of the wavelet are reported in both time  $|\psi_{a,b}(t)|^2$  and frequency  $|\widehat{\psi}_{a,b}(f)|^2$  domains. The Heisenberg boxes have side lengths  $2\sigma_t$  by  $2\sigma_f$  where  $\sigma_t$  and  $\sigma_f$  are respectively the standard deviations around the mean of  $|\psi_{a,b}(t)|^2$  and  $|\widehat{\psi}_{a,b}(f)|^2$ . These two standard deviations are used to quantify the spread around the mean of  $|\psi_{a,b}(t)|^2$  and  $|\widehat{\psi}_{a,b}(f)|^2$ , with large values corresponding respectively to reduced time and frequency precision.

Minimizing the area of the Heisenberg boxes we can find the best wavelet that maximises the time and frequency resolution. In the selection of the best wavelet this consideration is very important.

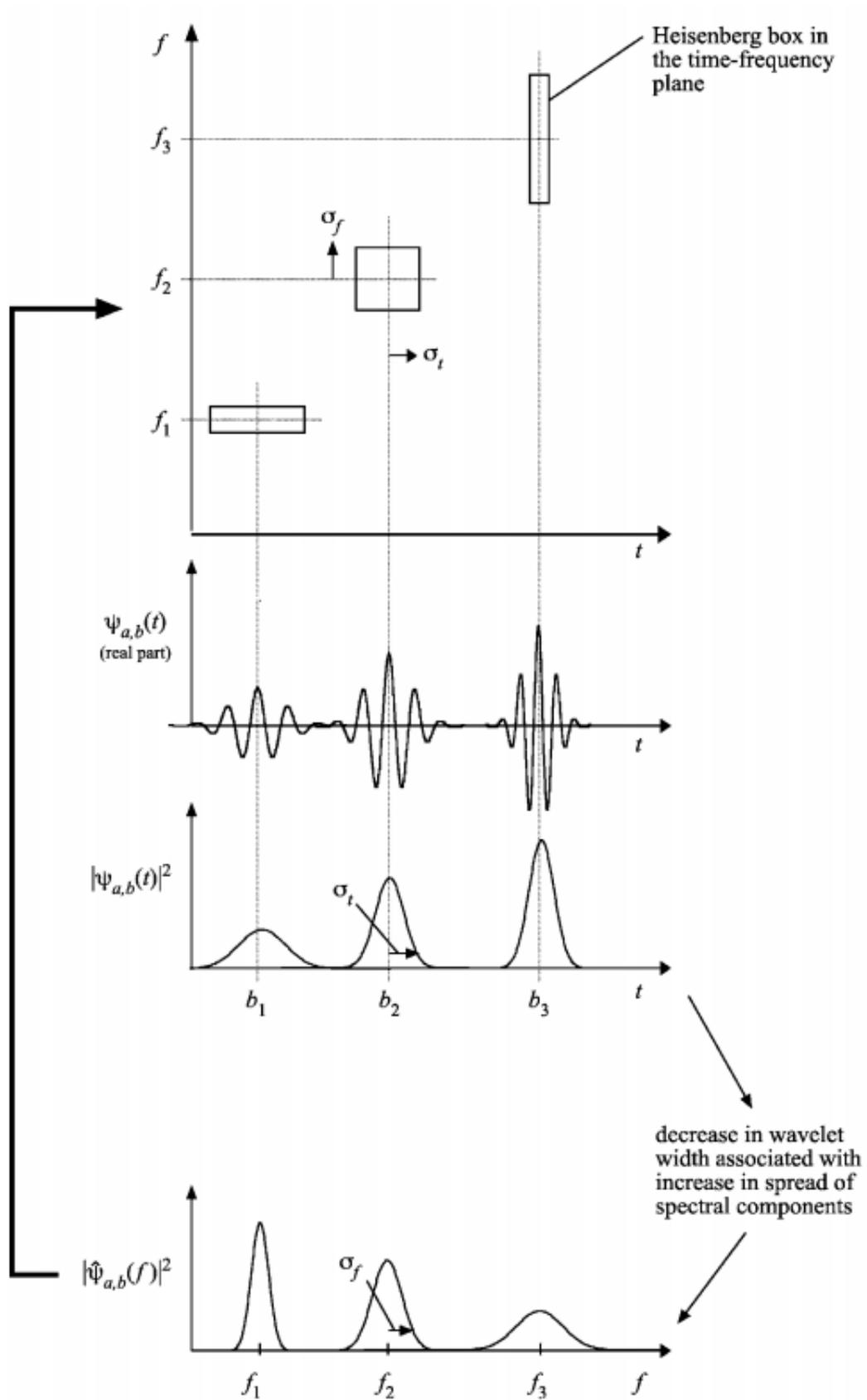


Figure 38: Heisenberg boxes in time-frequency domain. Heisenberg boxes at the top, Morlet wavelet real parts at three different scales and their energies in the middle-bottom. Source: (Addison 2002).

### 6.1.3 Morlet and Analytic Wavelets

The analytic wavelet transform (AWT) is a special family of complex-valued wavelet transforms (Lilly and Olhede 2010). As reported in Lilly and Olhede's work, these wavelets are particularly useful for the analysis of modulated oscillatory signals (Delprat et al. 1992, Mallat 1999, Carmona, Hwang and Torr sani 1997, Carmona, Hwang and Torr sani 1999, Scheper and Teolis 2003) and discontinuities (Tu, Hwang and Ho 2005). Details about all the properties and references for complex-valued wavelet transforms are reported by Selesnick's and colleagues (Selesnick, Baraniuk and Kingsbury 2005).

In this work we choose to use this family of wavelet transforms because the signal to examine is an electromyography (EMG) which is an oscillatory signal (Reaz, Hussain and Mohd-Yasin 2006).

The advantage of using complex-valued wavelets is that we can separate the amplitude and phase of the signal because the wavelet has real and imaginary parts. The phase information is more accurate than the modulus to reveal isolated singularities in a signal (Aldroubi and Unser 1996). The phase is used also to identify several types of transition points (i.e. local maxima) within the analysed signal (Haddad and Serdijn 2009).

It is important to remember that complex-valued wavelet transforms have Fourier transform equal to zero for negative frequencies as described in the third requirement to be a wavelet in section 6.1.

In this section two continuous wavelet transforms are discussed: Morlet wavelet and analytic Morse wavelet.

The Morlet mother wavelet is defined as:

$$\psi(t) = \pi^{-1/4} (e^{i2\pi f_0 t} - e^{-(2\pi f_0)^2/2}) e^{-t^2/2} \quad (6.9)$$

where  $f_0$  is the central frequency of the mother wavelet. The term  $e^{-(2\pi f_0)^2/2}$  is a correction term and it is used to correct the non-zero mean of the complex sinusoid described in the previous term. However if  $f_0 \gg 0$  the term  $e^{-(2\pi f_0)^2/2}$  becomes negligible and can be ignored writing the Morlet wavelet in the simple form:

$$\psi(t) = \frac{1}{\pi^{1/4}} e^{i2\pi f_0 t} e^{-t^2/2} \quad (6.10)$$

where  $\frac{1}{\pi^{1/4}}$  is the normalization factor,  $e^{i2\pi f_0 t}$  is a complex sinusoid and  $e^{-t^2/2}$  is a Gaussian envelope with a unit standard deviation. Then the Morlet wavelet for  $f_0 \gg 0$  is simply a complex wave ( $e^{i2\pi f_0 t}$ ) within a Gaussian envelope ( $e^{-t^2/2}$ ).

The dilated and translated Morlet wavelet could be defined substituting  $t$  with  $(t - b)/a$  in the expression 6.10:

$$\psi\left(\frac{t-b}{a}\right) = \frac{1}{\pi^{1/4}} e^{i2\pi f_0 [(t-b)/a]} e^{-[(t-b)/a]^2/2} \quad (6.11)$$

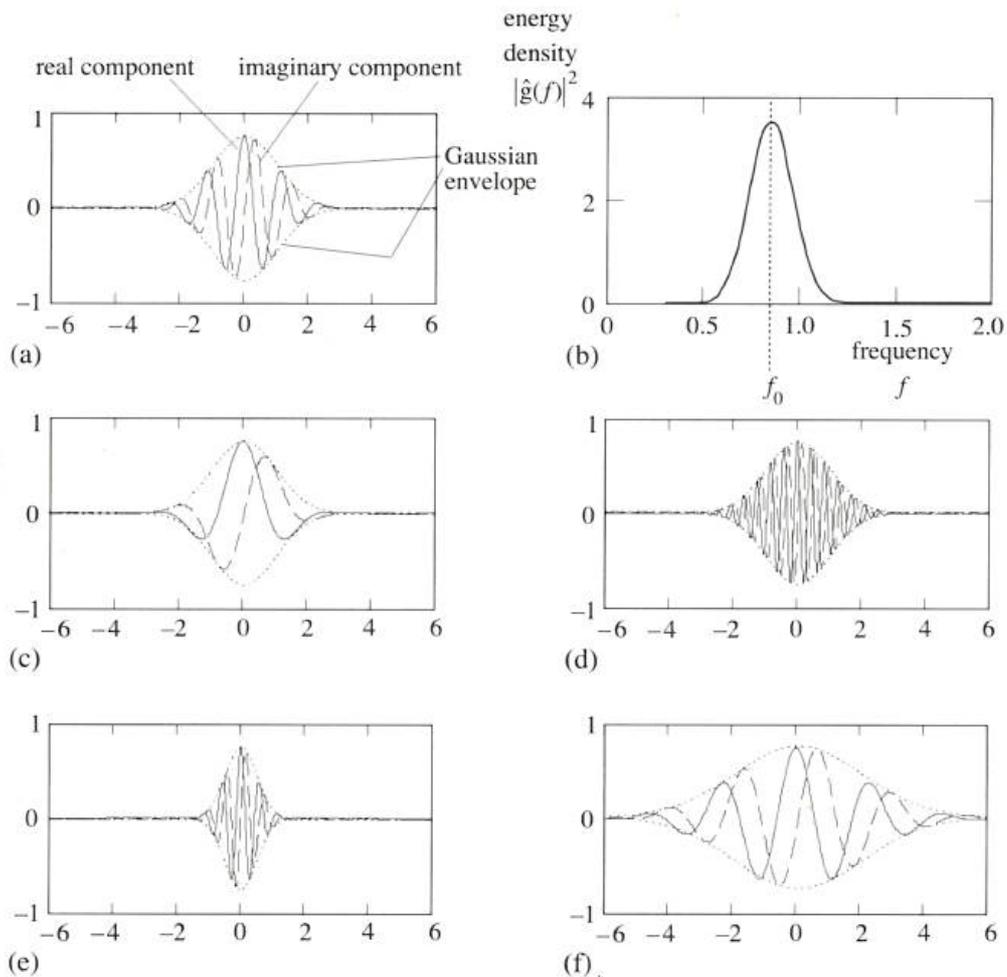
The imaginary and the real part of the mother wavelet described by equation 6.10 differ by a quarter of period. The normalization factor  $\frac{1}{\pi^{1/4}}$  is important to ensure that the wavelet has unit energy. It is important to notice that the wavelet described by equation 6.10 is not a real wavelet because it has non-zero mean equivalent to say that it has zero frequency component and then the second requirement described in section 6.1 is not strictly satisfied. However, if  $f_0 \gg 0$  equation 6.10 can be used with a minimal error.

The imaginary and real parts of the Morlet wavelet with its Gaussian envelope are shown in figure 39(a). This illustrates how the imaginary part is shifted by a quarter of period with respect to the real part.

The energy spectrum of the Morlet wavelet, given in figure 39(b), is defined as:

$$|\widehat{\psi}(t)|^2 = 2\pi^{1/2} e^{-(2\pi f - 2\pi f_0)^2} \quad (6.12)$$

where  $\widehat{\psi}(t)$  is the wavelet Fourier transform. The wavelet energy is given by integrating the equation 6.12. This energy is equal to one considering equation 6.11 ( $\pi^{-1/4}$  is the normalization factor).



**Figure 39. Morlet wavelet.**

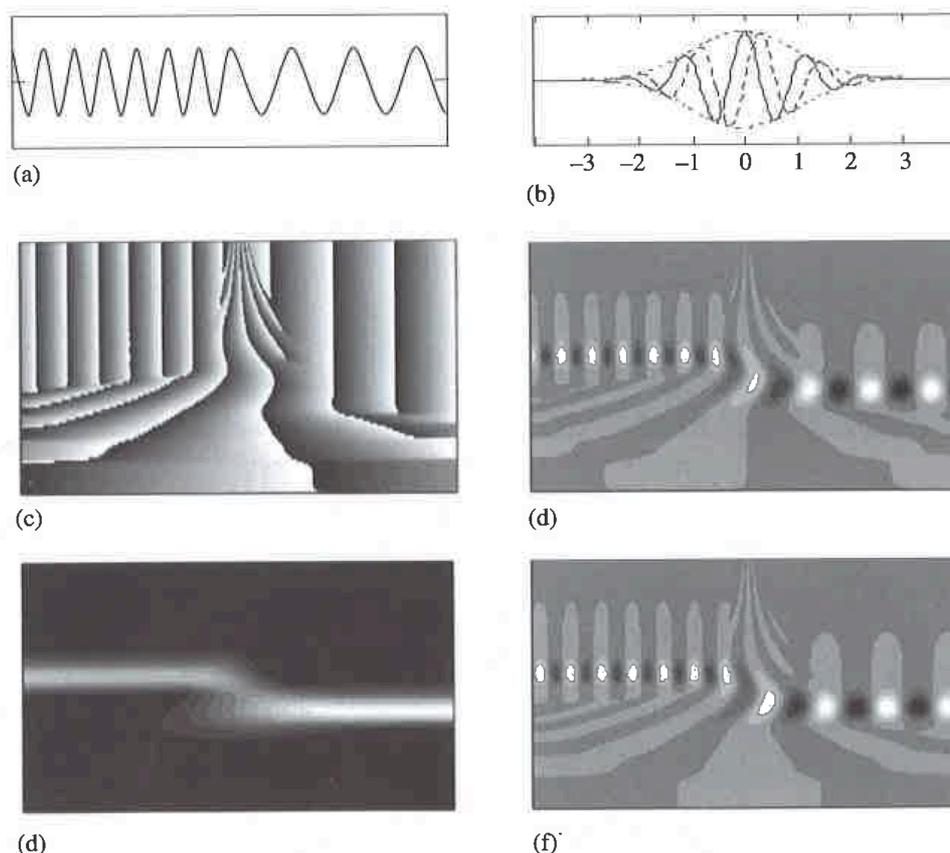
**a) Morlet wavelet with  $a = 1$  and  $f_0 = 0.894$ . b) Energy spectrum of the Morlet wavelet reported in a. c) Morlet wavelet with  $a = 1$  and  $f_0 = 0.318$ . d) Morlet wavelet with  $a = 1$  and  $f_0 = 1.909$ . e) Morlet wavelet with  $a = 0.5$  and  $f_0 = 0.894$ . f) Morlet wavelet with  $a = 2$  and  $f_0 = 0.894$ . Source (Addison et al. 2002)**

The central frequency  $f_0$  is the complex sinusoid frequency and determines how many oscillations are present in the Gaussian envelope. The effects of changing  $f_0$  are shown in figure 39(c) and 39(d). When  $f_0$  is smaller there are fewer oscillations (c), when it is bigger the number of oscillations is large (d). We can say that  $f_0$  is proportional to the number of oscillations contained in the Gaussian envelope. Usually  $f_0$  is 0.849 which corresponds of an angular frequency value  $\omega_0$  of 5.336 ( $\omega_0 = 2\pi f_0$ ). In practice  $\omega_0$  has a value between 5 and 6. If  $\omega_0$  is chosen less than five corresponding to  $f_0 < 0.8$ , equation 6.9 has to be used instead of equation 6.10 to compute the Morlet wavelet.

In figure 39(e) the squeezed Morlet wavelet ( $a = 0.5$ ) is shown, while the stretched Morlet wavelet is illustrated in figure 39(f) ( $a = 2$ ).

A signal with an abrupt change in the periodicity (a) and a Morlet wavelet (b) used to identify the abrupt change are represented in figure 40. The wavelet transform  $T(a, b)$  is a complex number and can be expressed as  $T(a, b) = \text{Re}(T(a, b)) + \text{Im}(T(a, b))$ . From this expression it is possible to find the wavelet transform phase

$\phi(a,b)$  illustrated in figure 40(c) and the wavelet transform modulus  $|T(a,b)|$  illustrated in figure 40(e). In figure 40(d) and 40(f) the real and the imaginary parts of the wavelet transform are shown respectively. Examining the figure it is possible to observe that they are very similar, in fact the imaginary waveform is simply a phase-shifted version of the real plot. The imaginary part of Morlet best matches the signal features present one quarter of cycle later than the real part. The change of periodicity for the signal is evidenced in all the wavelet plots of figure 40.



**Figure 40.** Use of the Morlet wavelet to detect an abrupt change in a signal periodicity. a) Signal to decompose b) Morlet wavelet with  $f_0 = 0.849$  and  $a = 1$ . c) Phase of the wavelet transform. d) Real part of the wavelet transform. e) Modulus of the wavelet transform. f) Imaginary part of the wavelet transform. Source (Addison et al. 2002)

Generalized Morse wavelets were introduced in Olhede's and colleagues (Olhede and Walden 2002). Lilly and Olhede developed Morse wavelet theory and its applications to modulated oscillatory signal analysis in their papers (Lilly and Olhede 2009, Lilly and Olhede 2010, Lilly and Olhede 2012).

The generalized Morse wavelet, according to Olhede and Lilly, has the following frequency-domain form:

$$\Psi_{\beta,\gamma}(\omega) = U(\omega)a_{\beta,\gamma}\omega^{\beta}e^{-\omega^{\gamma}} \quad (6.13)$$

where  $U(\omega)$  is the heaviside step function: a discontinuous function equal to zero for the negative values and equal to one for positive values (Bracewell 2000). The presence of  $U(\omega)$  in the expression is necessary to satisfy the third admissibility condition described in section 6.1. In equation 6.13  $a_{\beta,\gamma}$  is the normalizing constant defined as:

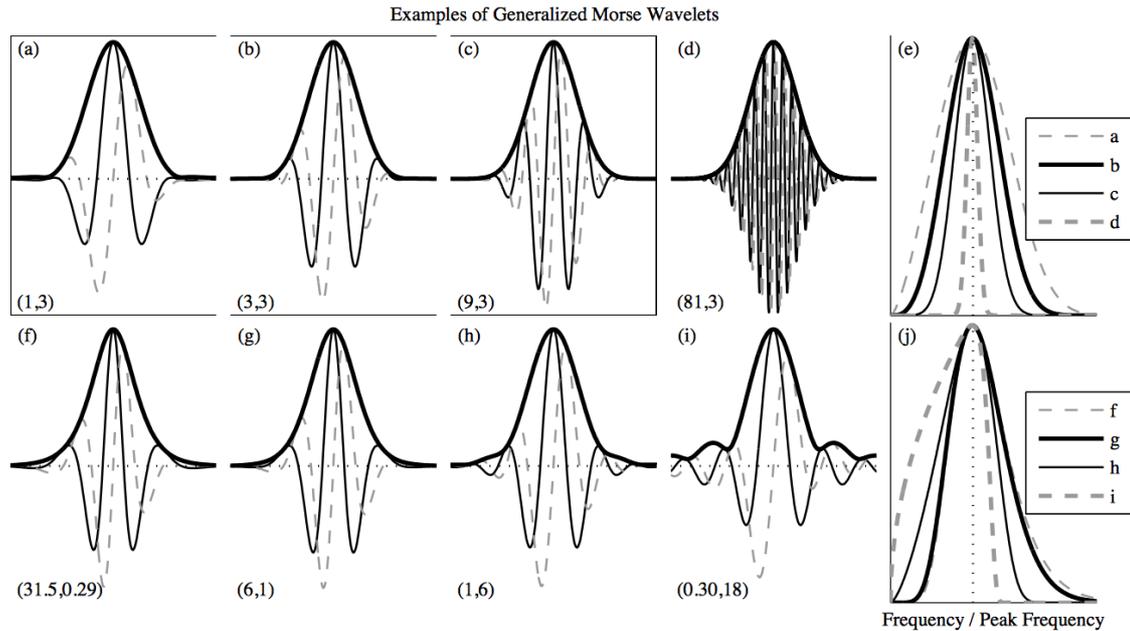
$$a_{\beta,\gamma} = 2 \left( \frac{e\gamma}{\beta} \right)^{\beta/\gamma} \quad (6.14)$$

Then the Morse wavelet shape is defined by the two parameters  $\beta$  and  $\gamma$ . The effects of changing these two parameters is examined in details in (Lilly and Olhede 2009).

Equation 6.13 is defined in the frequency domain. The inverse Fourier transform is used in order to derive the corresponding time domain equation:

$$\Psi_{\beta,\gamma}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_{\beta,\gamma}(\omega) e^{i\omega t} d\omega. \quad (6.15)$$

In figure 41 different generalized Morse wavelets are illustrated examining the differences in changing the two parameters  $\beta$  and  $\gamma$  (Lilly and Olhede 2010). The upper row, in figure 41, shows the effect of increasing  $\beta$  leaving constant  $\gamma = 3$ . In time domain increasing  $\beta$  makes the signal more oscillatory, while in frequency domain (figure 41(e)) makes the peak narrower. The lower row of figure 41 shows the effect of increasing  $\gamma$  and decreasing  $\beta$ . In this case, in the time domain the oscillations inside the central window do not change but the time behaviour of the wavelet changes. In frequency domain there is an enhancement to the right of the frequency peak shifts to the left of the frequency peak as  $\gamma$  increases.



**Figure 41. Examples of generalized Morse wavelets.**

**Panels (a–d) and (f–i) show wavelets in the time domain for different  $(\beta, \gamma)$  values, which are indicated in the lower left corner of each panel; for presentation, the wavelets are rescaled by their maximum amplitude. The real part is a solid line, the imaginary part is dashed, and the modulus is a thick solid line. The frequency-domain versions of the wavelets in the top and bottom rows are then given in panels (e) and (j) respectively. Source (Lilly and Olhede 2010)**

Figure 42 shows the area of Heisenberg boxes described in section 6.1.2 as a function of the  $\beta$  and  $\gamma$  parameters, both time and frequency axes have been rescaled for presentational clarity.

As described in section 6.1.2, it is important to minimize the area of the Heisenberg boxes to have the best resolution precision in the two dimensions. We decided to use  $\beta = 9$  and  $\gamma = 3$  because this corresponds to a small Heisenberg area giving us a good resolution precision in both dimensions (figure 42). The shape of the wavelet used is reported in figure 41(c). The letters contained in figure 42 indicate some well-known wavelet families: “L” for the lognormal wavelets; “C” for the Cauchy wavelets; “G” for the Derivative of Gaussian wavelets; “A” for the Airy wavelets; “e” for complex exponentials; “S” for the Shannon wavelet; “a” for the analytic filter; and “B” for the Bessel wavelet. Details about these wavelet families are reported in (Lilly and Olhede 2012).

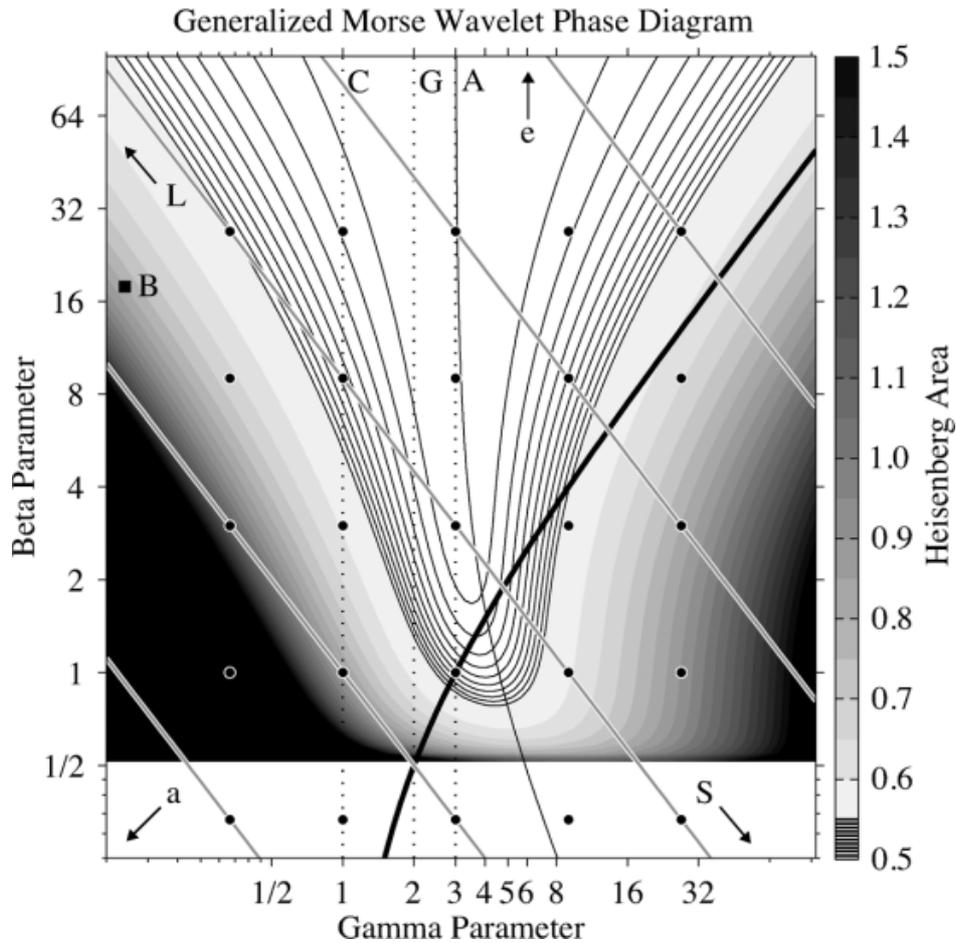


Figure 42. Parameter space for the generalized Morse wavelet superfamily. The Heisenberg areas are shaded over  $\beta$  and  $\gamma$  on log-log axes. Note that for  $\beta \leq \frac{1}{2}$  the Heisenberg area is undefined. Source: (Lilly and Olhede 2012).

In this work to pre-process the EMG (Reaz et al. 2006) we chose the generalized Morse wavelet because it is more versatile than Morlet wavelet. In fact Morse wavelet uses two parameters instead of only one, allowing a better localization in frequency or time (Aguar - Conraria and Soares 2014). As Lilly and Olhede illustrated in their work: “With two free parameters, the generalized Morse wavelets can take on a broad range of forms which has not yet been fully explored, and in fact this family encompasses most other popular analytic wavelets”(Lilly and Olhede 2009).

## 6.2 THE DISCRETE WAVELET TRANSFORM

In this section a brief introduction of discrete wavelet transform is given, for more details refer to (Addison 2002).

The discrete wavelet transform differs from the continuous wavelet transform, reported in the equation 6.5 in the fact that it uses discrete values for the dilation and translation parameters  $a$  and  $b$ . The logarithmic discretization can be used to

discretize  $a$  and then link this to the size of the steps between locations of translation.

The idea is to move with discrete steps proportional to the  $a$  scale between the different  $b$  locations. Then the wavelet has the following form:

$$\Psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \Psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \quad (6.16)$$

where  $w(a) = \frac{1}{\sqrt{a_0^m}}$ ,  $m$  and  $n$  are integers used to control the dilation and the translation,  $a_0$  is a fixed dilation step parameter greater than one and  $b_0$  is the location parameter which must be greater than zero. The size of the translation step is  $\Delta b = b_0 a_0^m$  is directly proportional to the scale parameter  $a_0^m$ .

The discrete wavelet transform of a continuous signal  $x(t)$  is then:

$$T_{m,n} = \int_{-\infty}^{\infty} x(t) = \frac{1}{a_0^{m/2}} \Psi(a_0^{-m}t - nb_0) \quad (6.17)$$

where  $T_{m,n}$  is called the wavelet transform or detail coefficients gives the values on a scale-location grid with indexes  $m$  and  $n$ .

While CWT operates at every scale giving as a result a signal larger than the original one, DWT decomposes the signal into approximation low frequency and detail high frequency coefficients, giving as a result a signal of the same length as the original (Nguï et al. 2013). CWT is scaled and shifted over the full domain of the analysed signal, while DWT is scaled and shifted only on a subset of the signal domain determined by the discrete steps.

DWT are widely used for image processing (Broughton 1998), signal processing of accelerations for gait analysis (Martin 2011) and in many others areas (Akansu et al. 1998, Akansu and Medley 2006, Akansu and Smith 2012).

DWT are not used here because, as explained previously, it is computed only on a domain subset of the signal as opposite to CWT that is computed across the entire signal domain. We preferred CWT to pre-process EMG signal across its entire domain avoiding the loss of some useful information.

### 6.3 SUMMARY AND CONCLUSIONS

In this chapter the wavelet transform is discussed. The following key-points are highlighted:

- Wavelet transform is used for time-frequency signal analysis. Time-frequency analysis shows all the frequency is contained in the signal along with the time at which they occur.
- A wavelet function has a wave-like form that respect defined mathematical constraints.
- The two main operations on the wavelet function are: scaling and translation. With scaling the wavelet is stretched/squeezed along the signal, while with translation it is moved along the time domain of the signal.
- The wavelet transform is the convolution of the wavelet function with the signal considered. It transforms the time domain signal in a two-dimensional domain signal considering frequency (inverse of scaling) and location.
- There are two kinds of wavelet transform: continuous wavelet transform (CWT) where the wavelet transform is computed along the entire signal domain and the discrete wavelet transform (DWT) where the wavelet transform is shifted and scaled on a subset of the signal domain determined by the discrete steps.
- In this work CWT is used to pre-process EMG signal since it is more appropriate to compute the transformation along the entire signal domain.
- The generalized Morse wavelet transform is used in this work. Generalized Morse wavelet is an analytic wavelet, a family of complex-valued wavelets useful to separate amplitude and phase of the signal having real and imaginary part. These wavelets are versatile thanks to the two parameters  $\beta$  and  $\gamma$ , which regulate their shape.

In conclusion the wavelet transform is preferred to the Fourier transform because it permits to know not only the frequencies contained in the signal but also the time in which they occur. Morse wavelet is used instead than the well-known Morlet wavelet because it is more versatile using two parameters instead of one. The two parameters of the wavelets are chosen in order to minimize the Heisenberg boxes area improving precision in both time and frequency domain. The main idea of this work is using the wavelet transform to pre-process the EMG signals in order to find useful features able to discriminate among healthy subjects, dystonia patients and subsets of dystonia patients. In the next chapter the hand-opening experiment is described, with details about the pre-processing of the EMG using wavelets and the feature extracted.

# CHAPTER 7: HAND OPENING-CLOSING EXPERIMENT

In this chapter the second part of the thesis work is described. Several dystonia patients and healthy subjects were measured while performing different tasks. The data recorded are the electromyography (EMG) of four different muscles. EMG data are pre-processed with wavelet transform described in Chapter 6.

The aim of this second part is to distinguish between dystonia patients and healthy subjects but also between functional and organic dystonia patients (section 4.4).

The chapter is organised in five main sections. Section 7.1 reports the aim and motivations of the work reviewing the previous studies. Section 7.2 gives an introduction about the EMG signal explaining how it is computed and describing the two different kinds of EMG (surface and intramuscular). Section 7.3 reports the methodology used, describing: the subjects recruited (section 7.3.1), the equipment used (section 7.3.2), the experiment (section 7.3.3) and the data used along with the pre-processing (section 7.3.4). Section 7.4 reports the results divided for the different inputs considered plus a subsection (section 7.4.5) showing a diagram of a classifier evolved with the derived mathematical expression. Section 7.5 reports the conclusions with some suggestions for future works.

## 7.1 AIM AND MOTIVATIONS

In the experiment described, two different subgroups of dystonia patients and a group of healthy subject perform several movements involving upper limbs. During the experiment the following pairs of EMG were recorded: biceps/triceps and flexor digitorum superficialis/extensor digitorum communis (FDS/EDC). The patients were diagnosed with two kinds of dystonia: organic dystonia (comprising cervical dystonia, focal and secondary dystonia involving upper limb muscles) and functional dystonia.

The aim of this study is using these EMG recordings to distinguish among dystonia patients and healthy subjects but also between the two big subgroups of patients: organic and functional (section 4.4). Clearly the muscles affected by the disease are determined by the different kind of dystonia, although there is evidence in the literature that patients present abnormal muscle activities also in non-dystonic body parts (De Vries et al. 2007). De Vries and colleagues measured the activity of two wrist muscles (flexor and extensor) of eight cervical dystonia patients and eight healthy subjects during a flexion-extension movement of the right wrist. The results showed abnormal muscle activities for the patients, evidenced by lower mean EMG amplitude and a prolonged extensor muscle contraction compared with healthy subjects. The results support the view that “although dystonic involvement is clinically only seen in one part of the body, other parts of the body may exhibit subclinical dystonic movement abnormalities” (De Vries et al. 2007). The muscles

exhibiting abnormal muscle activity are in “pre-dystonic state”. We used these findings including in our study different kinds of dystonia such as cervical dystonia that do not affect directly the upper limb muscles. The first step is distinguishing between the dystonia patients and healthy subjects verifying the possible abnormal activity also in the muscles not directly affected by dystonia.

The big challenge of this study is distinguishing between organic and functional patients. Although this classification could be very challenging, it could lead to a discovery of some criteria useful to diagnose functional dystonia. In the literature there is no clear agreement on diagnostic criteria for functional dystonia (Morgante et al. 2012, Ganos et al. 2014). Historical variables and/or psychiatric comorbidity (co-occurrence of one or more diseases or disorders in an individual) which have poor predictive values (Espay and Lang 2015, Pareés et al. 2014) are usually used for the diagnosis. The need of clear diagnostic criteria based on laboratory studies has been highlighted (Espay and Lang 2015). These criteria should permit earlier diagnosis and treatment reducing the risk of long-term disability and reducing at the same time the health and social care costs (Birmingham et al. 2010).

The diagnosis of organic dystonia is based on its core motor features and temporal evolution (Albanese et al. 2013). There are also some typical non-motor features including cognitive and psychiatric features (Stamelou et al. 2011). However there are some atypical forms of organic dystonia which have been a challenge to diagnose even for expert clinicians (Pandey et al. 2014, Bentivoglio et al. 2002). For this reason a misdiagnosis rate of 25-52% is currently experienced (Pal 2011).

The diagnosis of functional dystonia is based on assessment of inconsistency and incongruence (in both the history and neurological examination) with organic disease patterns (Espay and Lang 2015, Ganos et al. 2014). However there may be a significant number of functional patients overlapping with the atypical form of dystonia making it difficult the diagnosis.

All these considerations evidence the need of more precise criteria to distinguish between organic and functional dystonia. From literature findings we know that functional and organic dystonia share some features such as co-contraction (Pal 2011, Macerollo et al. 2015) and reduced intra-spinal and intra-cortical inhibition (Espay et al. 2006). Potentially discriminating features for functional dystonia are: greater bradykinesia (Criswell et al. 2010), co-activation of non-contiguous muscles (Mehta et al. 2013), less agonist/antagonist co-contraction and faster reaction times (Macerollo et al. 2015). However the utility of these findings is limited: co-contraction is not ubiquitous in organic dystonia (so absence/reduction is not specific for functional dystonia) (Malfait and Sanger 2007) and intra-cortical inhibition and reaction times are influenced by such factors as attention to the limb and personality disorder (Ganos et al. 2014). Also functional dystonia is challenging to study experimentally due to the difficulties to diagnose the pathology itself. Then the studies on the argument are few and only a small sample size of patients is considered (Macerollo et al. 2015, Mehta et al. 2013).

The approach used in this study to improve the classification of organic and functional dystonia patients involves the use of evolutionary algorithms described in Chapter 2 to analyse the possible differences between the subgroups of patients. The main idea is helping doctor's diagnosis through the identification of some discriminating features.

## **7.2 EMG SIGNAL**

The Electromyography is a way to measure the muscle activity considering the electrical signal emanated by the muscle.

Usually what is studied is the voluntary neuromuscular activation of muscles during postural tasks, specific movements, work conditions and treatment/training regimes. Konrad gave a clear explanation of muscle contraction process (Konrad 2006). His work is summarized in this section explaining the process of muscles contraction and how EMG detects muscle activity.

The "motor unit" represents the smallest functional unit to describe the neural control of the muscular contraction. As Konrad explained: "motor unit is the union of the cell body and dendrites of a motor neuron, the multiple branches of its axon, and the muscle fibers that innervates it" (Konrad 2006).

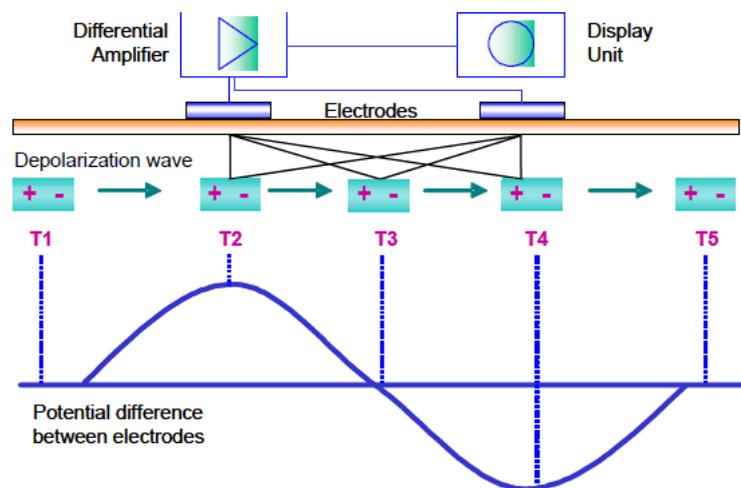
The resting potential at the muscle fiber membrane, which is approximately between -80 to -90 mV (when the cell is not contracted), is due to the ionic difference between the inner and outer spaces of a muscle cell. When the muscle is not contracted, the intracellular space has a negative charge compared to the external surface. In order to start a muscle contraction, the central nervous system or a reflex induces the activation of an alpha-motor neuron, which is a particular neuron responsible for initiating the contraction of a muscle. The activation of an alpha-motor neuron results in the conduction of the excitation along the motor nerve, which is a particular nerve responsible to carry out of the central nervous system the command information and send it to the muscles. After the command information reaches the muscle, there is the formation of a potential at the muscle fiber innervated by its motor unit. The fiber membrane is depolarized; the cell becomes less negative than before. The entrance of sodium ions inside the cell causes the depolarization. The repolarization is activated then and restores the negative charge of the cell through a backward exchange of ions within the active ion pump mechanism. If the level of sodium ions exceeds a certain threshold, the depolarization causes a quickly change in membrane potential from - 80 mV up to + 30 mV. The quick change of membrane potential is a phenomenon called action potential. The repolarization immediately restores this electrical burst.

The action potential starts at the extremity of the muscles and propagates inside and along the muscle. As a consequence of the excitation, calcium ions are released inside the intra-cellular space. The chemical processes derived are the cause of shortening of the contractile elements of the muscle cell resulting in the contraction of the muscle.

The EMG signal measures the action potentials at the muscle membrane resulting from the depolarization and the repolarization process described.

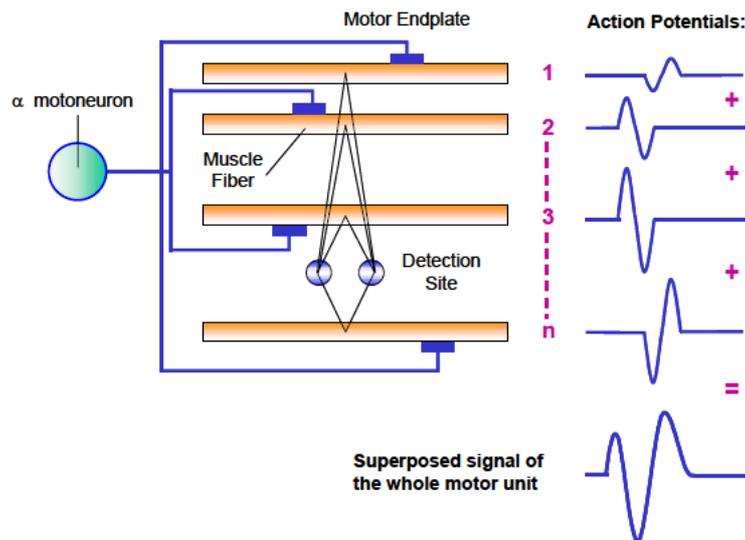
The depolarization–repolarization cycle results in a depolarization wave or electrical dipole, which travels along the surface of the muscle fiber.

Usually EMG measures are done through bipolar electrode configurations and a differential amplification. Figure 43 shows a detection of a single muscle fiber. The dipole generates a potential difference between the electrodes dependant on the spatial distance between electrode 1 and 2.



**Figure 43. Model of the single muscle fiber detection by using a differential measure of muscle activity.  
Source: (Konrad 2006)**

The model illustrated in figure 43, shows how the monopolar action potential creates a bipolar signal within the differential amplification process. This is due to the fact that the electrode pair can see the magnitude of all innervated fibers contained in a motor unit. Figure 44 shows how all the magnitudes of all innervated fibers sum up to a triphasic motor unit action potential with shape and size depending on the geometrical fiber orientation in ratio to the electrode site (Konrad 2006).



**Figure 44. Model of the triphasic motor unit action potential.**  
 Source: (Konrad 2006)

Electromyography (EMG) signal is then a biomedical signal measuring the neuromuscular activities. EMG measures precisely the electrical currents generated in the muscles when they contract/relax receiving the command by the nervous system (Reaz et al. 2006). This is done measuring the electric potential produced by muscle cells when these are electrically or neurologically activated with the process described previously.

The instrument used to measure the EMG is called electromyograph and the record produced is an electromyogram.

It is possible to measure two kinds of EMG: surface EMG and intramuscular EMG. Surface EMG measures the muscle activity from the skin. They are placed on the skin, where the muscle is located. The skin is cleaned with an alcohol pad then the needle electrode is placed at the belly of the muscle: the longitudinal midline (Delsys technical note 101). More than one electrode is used because the measurement is expressed as the potential difference (voltage difference) between the electrodes. Usually a pair of electrodes is used for each muscle, but it is possible to use also a complex array of multiple electrodes.

The surface EMG placement usually is difficult and depends on the muscle chosen and its size. Also it has three main limitations: the skin is a possible source of interference and the more body fat the subject has the more EMG signals are weak; recordings can be done only on superficial muscles and it is difficult to distinguish the activity among adjacent muscles.

Intramuscular EMG is measured using needles or wires inserted directly into muscles (Merletti and Farina 2009). This technique is invasive but it overcomes the limitations of the surface EMG. The skin does not influence the signal, it is possible to record not only superficial muscles and also the needle is placed directly into the muscle so it is easier to distinguish among the activity of adjacent muscles.

From this discussion, everything seems to suggest that the intramuscular EMG is the correct choice. However this technique is very invasive and in our work we chose to use the surface EMG because one of the aims of this work is to design a non-invasive and easy test for the patients.

## **7.3 METHODOLOGY**

### **7.3.1 Subjects recruited**

Participants with dystonia were recruited from the current caseloads of the movement disorders consultants at Monash Medical Centre (MMC) in Melbourne (Australia). Three different classes of subjects were measured:

- 30 patients with organic dystonia
- 10 patients with functional dystonia
- 30 age-matched healthy subjects (controls)

Patients and controls subjects were chosen following the three inclusion criteria:

- Organic dystonia: Expert diagnosis, according to accepted guidelines, with upper limb or cervical involvement (genetic, idiopathic focal or secondary);
- Functional Dystonia: Expert diagnosis, documented or clinically established according to the Fahn-Marsden criteria (Burke et al. 1985), with upper limb involvement;
- Controls: Capacity to consent and able to perform assessments.

Subjects were removed from the study according the following exclusion criteria:

- Aged under 18 years;
- Lacking capacity to consent;
- Unable to perform movement assessments (e.g. due to cognitive deficit).

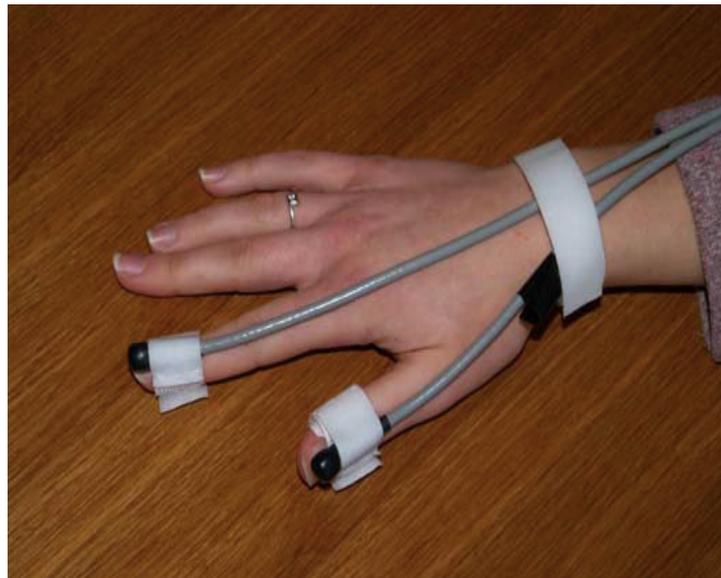
Control subjects were recruited from spouses and friends of the patients who attend clinics at MMC.

The assessments of the symptoms and of the disease were done according to the Fahn-Marsden (Burke et al. 1985) and Hinson (Hinson et al. 2005) rating scales. Scores were compared for inter-rater reliability. The mean of the rates' scores were used to correlate the movement data with diagnosis and symptoms severity. The ethical approval was obtained for the study and the protocol was clarified to subjects with a comprehensive information sheet, given to the subjects one week before the assessment. The same information sheet was presented again the day in which the subject signed the informed consent and the clinical investigator (Dr Rachel Newby) addressed any possible queries. Also the subjects were aware of the possibility to withdraw from the study at any time without any impact on their treatments.

### 7.3.2 Equipment used

The subjects performed different movements described in the next subsection. Two different kinds of data were recorded: the kinematic data and the muscular data.

The kinematic data were recorded using two Electromagnetic (EM) sensors attached to thumb and index fingers (figure 45). These two sensors are part of Polhemus Patriot an EM tracking device described in section 5.2.2 and shown in figure 23 and figure 24.

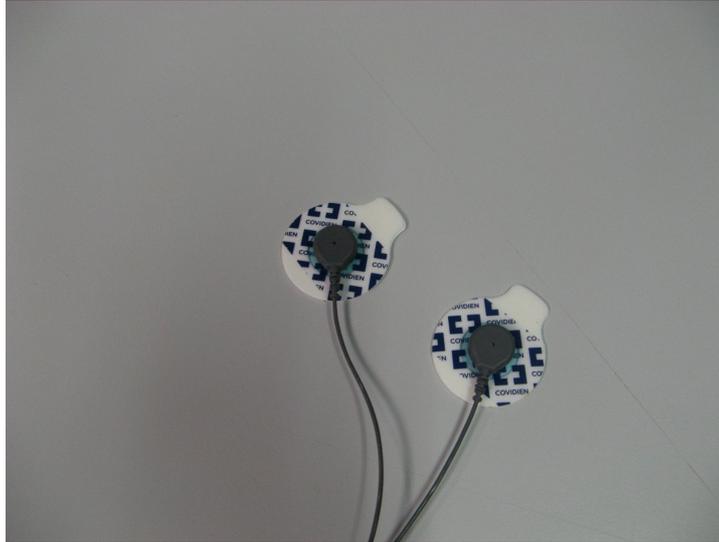


**Figure 45. Electromagnetic sensor attached to thumb and index fingers.**

**Picture produced by Dr Rachel Newby.**

The muscular data are measured using surface EMG (Shimmer <sup>TM</sup>) on the skin overlying four muscles of the upper limb. The muscles recorded are the following pairs: biceps/triceps and flexor digitorum superficialis/extensor digitorum communis (FDS/EDS).

Surface EMG electrodes shown in figure 46 are attached to boxes called Shimmer EMG units that collect and store the data (figure 47). Each shimmer EMG unit has two EMG channels to record the pair muscles, then each subject wore 8 EMG sensors (electrodes) attached to four Shimmer EMG units (left biceps/triceps, left FDS/EDC, right biceps/triceps, right FDS/EDC).



**Figure 46.** Shimmer EMG sensors fixed on the skin in correspondence of the selected muscles.  
Source: (Shimmer3 ECG Unit 2018).



**Figure 47.** Shimmer EMG unit secured to the arms with elasticated straps during the movements.  
Source: (Zebra press 2014).

### **7.3.3 Experiment description**

The experiment comprised many movements involving upper limbs, wearing the EMG sensors (figure 46-47) and the EM sensors (figure 45) described in the previous subsection. Each subject did the following tasks with each arm/hand (dominant and non-dominant):

1. Wrist flexion/extension ten times
2. Finger tapping for 15s x 3
3. Finger tapping in time with metronome x 3 (1, 2 and 3 Hz)
4. Hand opening and closing ten times x 2

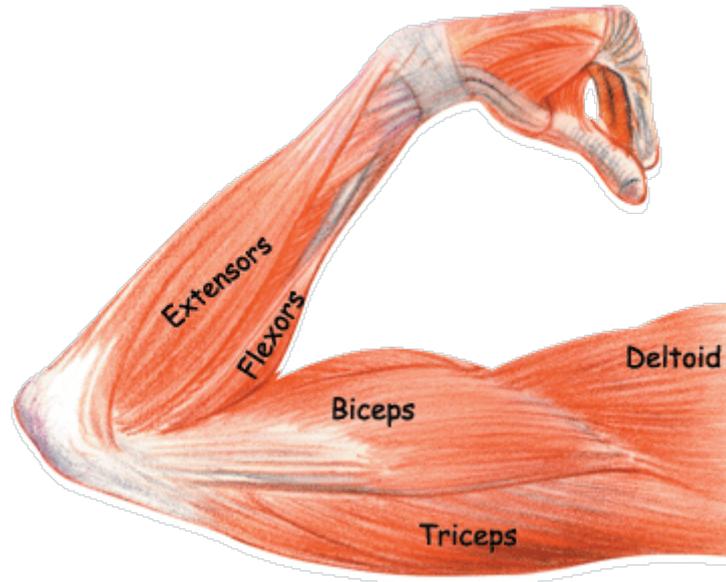
5. Arm pronation/supination ten times x 2
6. Resting (30s)
7. Arms outstretched (25s) x 2
8. Finger-nose pointing ten times x 2

The experiment also included some writing tasks using a digitising graphic pad and wearing only the EM sensor, however these tasks were not included in the scope of this study so were not taken in consideration for these experiments.

The EMG recording was done continuously for all the movements creating very large files. The first challenge was to find a way to separate the different movements. During the experiments the logs containing the start and the stop Unix time (The Open group 2017) for each task, were registered. We compared the start and stop of each task containing in the logs with the UNIX timestamp recorded from the shimmers during the experiment. This process permitted us to divide the large recording file into the eight small EMG recording files corresponding to each single task.

Having all the different tasks, the next challenge was to decide which of them was better in order to distinguish the different class groups (organic dystonia, functional dystonia and controls). We decided to choose a repetitive task in order to compute the wavelet transform on each trial of the task. In the task chosen the muscles measured had to be active as much as possible. The muscles measured, as described in the previous subsection, are: biceps, triceps, flexors (FDC) and extensors (EDC). They are located on the arm as shown in figure 48, so the finger tapping was excluded because the muscles on the arm are not very active in these movements. The arms outstretched task was excluded because it is not a repetitive movement. So four different tasks remained for the choice: wrist flexion/extension, hand opening/closing, arm pronation/supination and finger-nose pointing. The hand opening/closing task was chosen because it was easy to determine the single trials during the task using the kinematic data. In particular the distance between the index finger and the thumb was used (details in the next subsection). In the other tasks we did not find an appropriate way to compute the single trials using the data available.

Then from this point all the data analysis and the results will be referred only to the hand opening-closing task.



**Figure 48. Location of the muscles on the arm.**  
 Source: (workout trends 2013)

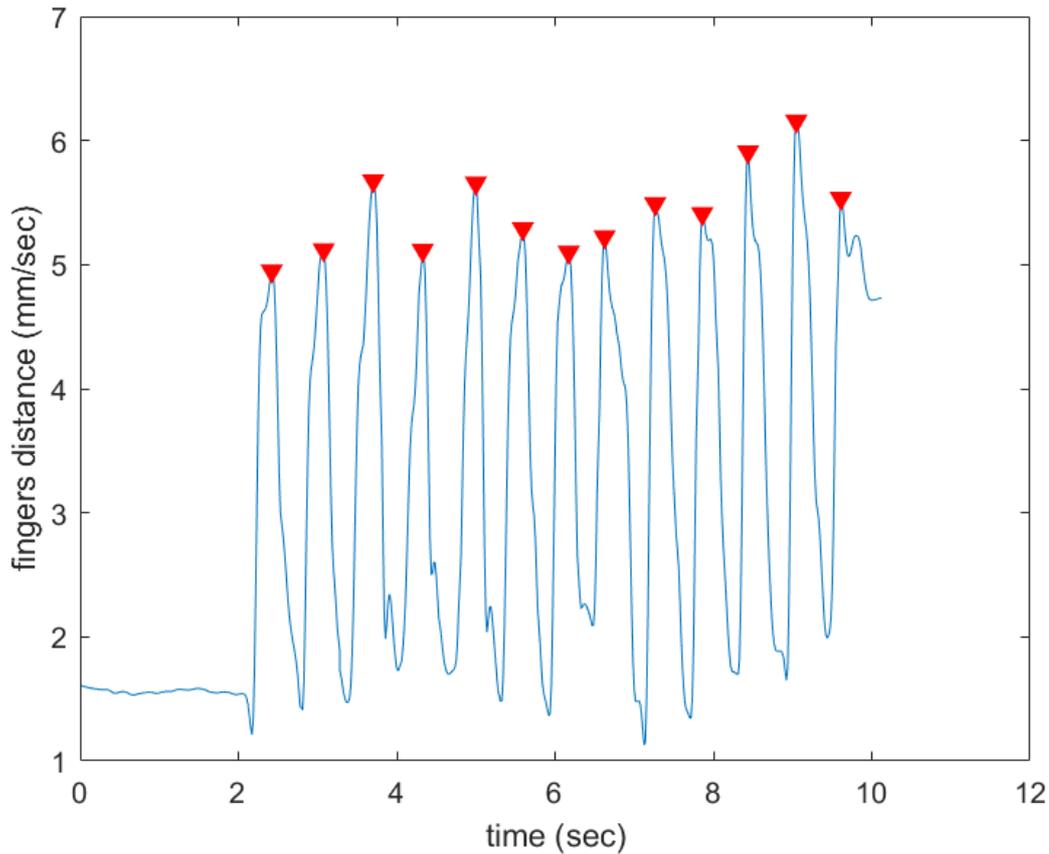
### 7.3.4 Hand opening-closing data

First of all the EMG data were filtered with a high pass filter with cut-off frequency of 10 Hz. This was done to remove the artifacts in the movement. These artifacts are due to different source of noise such as interferences, other electromagnetic signal and so on.

After filtering the EMG data the kinematic data were analysed. The kinematic data were filtered too with a low pass Butterworth filter (Zhongshen 2003) at 5Hz to remove extreme data. Then the Euclidean distance between the index and the thumb fingers is computed as following:

$$\text{dist} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (7.1)$$

where  $x_1, y_1, z_1$  and  $x_2, y_2, z_2$  are respectively the Cartesian coordinates recorded from the two EM sensors on the thumb and the index finger. In figure 49 an example of distance between the fingers relative to a patient is reported. The distance reported on y-axis is computed with (7.1), while the time on x-axis is computed using the UNIX timestamp retrieved from the EM sensors. In the figure, the peaks are highlighted in red and correspond to the maximum distance between the fingers (when the hand is open). The peaks were used as trigger points delimiting the single trials during the hand opening-closing movement. In other words, the time of the trigger points was computed using the UNIX timestamps of the EM sensors and then the same timestamps were found in the timestamp recorded from the EMG sensor in order to determine the trigger points in the EMG data.

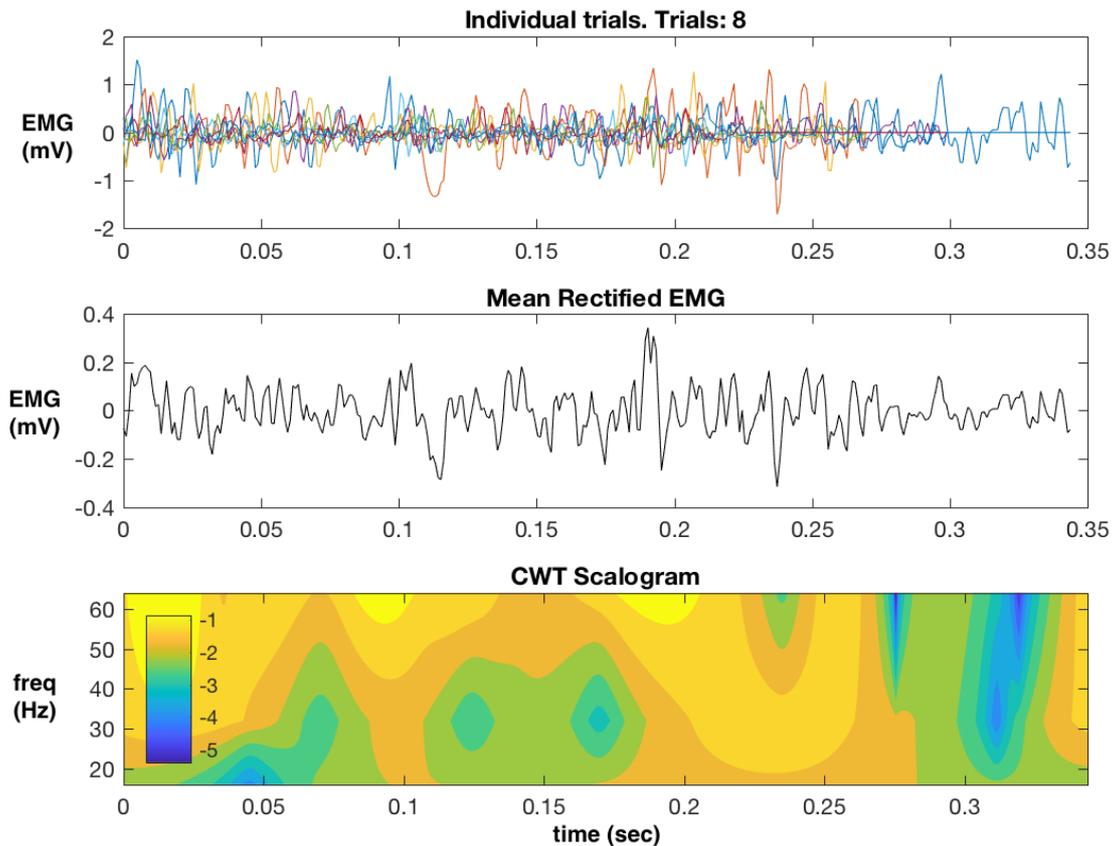


**Figure 49. Thumb-index fingers distance on y-axis, time on x-axis. The different peaks corresponding to the different hand-opening trial (maximum distance) are highlighted.**

The EMG data were pre-processed with CWT wavelet transform choosing in particular the generalized Morse wavelet transform ( $\gamma = 3$  and  $\beta = 9$ ) described in section 6.1.3. The range of frequencies included for the analysis is 16-32-64 Hz. This range of frequencies is considered enough to analyse human movements as referred in the literature.

The EMG data were divided in single trials using the location of the trigger points shown in figure 49. The Morse wavelet transform were computed on each single trial and then the average considering all trials was computed. All trials were represented as rows of a matrix representing all the data. In order to compose the rows of the matrix, the fact that the trials have different lengths had to be taken in consideration. All the trials then were zero-padded in order to reach the length of the maximum one.

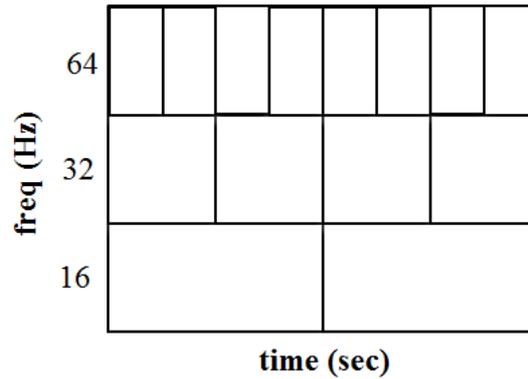
An example of EMG data and corresponding CWT spectrogram (section 6.1.1) is reported in figure 50.



**Figure 50. EMG data of a patient with the corresponding CWT scalogram.**  
**Top: EMG data for each different trial, middle: rectified EMG obtained as mean on the trials,**  
**bottom: CWT scalogram**

In figure 50, at the top, the different EMG trials are shown highlighting the zero-padding on some trials (e.g. blue one). The rectified EMG obtained as the mean along all the trials is reported in the middle of the figure. The wavelet transform spectrogram, reported at the bottom of the figure, reveals how much energy is present during the movement at a certain frequency (on y-axis) and time (on x-axis). The heat map on the left matches the colours with the energy power: yellow regions are where there is the maximum energy; blue area where there is the minimum.

We divided the spectrogram in different regions creating a grid. For the first frequency 16Hz we defined two regions, for the second frequency 32Hz four regions and for the third frequency 64 Hz eight regions. The grid created is shown in figure 51, with all the regions equispaced in time.



**Figure 51. Grid built on the spectrogram.**  
The spectrogram is divided in regions according to the different frequencies.

The grid built is useful to understand the differences among the classes in the different regions. The sub-divisions of the grid were chosen to represent the behaviour of the signal at different frequencies. For low frequencies the spectrogram does not change fast, so we chose to divide the 16Hz frequency in only two regions defined enough to represent the signal at 16Hz. The number of regions is incremented to four at 32 Hz when the spectrum changes faster and then the values considered change in the different regions. For the same reason the number of regions at 64 Hz become eight when the spectrum changes faster than before and then the values considered change too. The two main novelties of this work are: the use of the EMG wavelet scalogram and the construction on it of the grid reported in figure 51. To our knowledge there are not previous work that used the EMG wavelet scalogram and its division in regions in order to classify the different dystonia patients. For each region the mean, the max and the standard deviation were extracted and given as input to the SVM (Durgesh and Lekha 2010), ANN (Maind and Wankar 2014) and CGP (described in Chapter 2) to generate classifiers distinguishing among the different classes. The results obtained are reported in the next section.

## 7.4 RESULTS

The muscles measured during the experiment are four: biceps/triceps and FDS/EDC. Examining the EMG graphs during the hand opening-closing movement, we noticed that the biceps/triceps were not so active so we decided to include in our analysis only the EMG data relative to FDS/EDC.

In table 9 the number of subjects and samples for each muscles and class are reported. The numbers are less than expected because some data were discarded for recording errors or lack of consent.

Muscles	Subjects			Samples		
	Org	Fun	Con	Org	Fun	Con
FDS	25	7	23	72	16	64
EDC	24	6	25	76	11	75

**Table 9.**Number of subjects and samples considered for each class (organic, functional, controls) considering the muscles FDS and EDC.

In table 9 it is clear that the dataset is unbalanced in the fact that the functional samples are very few in comparison to the other classes. He and Garcia wrote a review about the possible techniques to deal with unbalanced data (He and Garcia 2009). Two possible sampling approaches are considered here: the oversampling and undersampling.

With the first approach the samples of the smallest class are increased creating new reasonable synthetic data from the existing samples. The method used for the oversampling was Adaptive Synthetic Sampling (ADA-SYN)(Haibo He et al. 2008). In this method a density distribution function is used to decide the number of synthetic samples that need to be generated for each sample  $x_i$  belonging to the minority class  $S_{min}$ . The density distribution function  $\tau_i$  is defined as following:

$$\tau_i = \frac{\Delta_i/K}{Z} \quad i = 1, \dots, |S_{min}| \quad (7.2)$$

Where  $\Delta_i$  is the number of samples in the K-nearest neighbors of  $x_i$  belonging to the majority class and Z is the normalization constant so that  $\sum \tau_i = 1$ . He and Garcia defined The K-nearest neighbors for each  $x_i \in S_{min}$ , for a specific integer K as: “the K elements of  $S_{min}$  whose euclidian distance between itself and  $x_i$  under consideration exhibits the smallest magnitude along the n-dimensions of feature space X” (He and Garcia 2009). The new random synthetic sample is created by selecting a random K-nearest neighbour of  $x_i$ , then multiplying its feature vector for a random number in [0,1] and finally adding the resulting vector to the original  $x_i$ . Results obtained using ADA-SYN were not so different from those using the original unbalanced data, so it was decided to discard this approach because the creation of valid synthetic data introduces significant uncertainty.

In the second approach the samples of the largest class are reduced with undersampling. The undersampling can be random or informed trying to minimize the information loss. We did not consider this approach because the reduction of the total dataset makes it impossible to have a number of sufficient samples in test, training and validation sets of each cross-validation fold.

We decided to continue with the data available taking into account that in some studies the classifiers induced with unbalanced dataset gave performance comparable to the ones induced with the respective balanced dataset computed with different techniques (Batista, Prati and Monard 2004, Japkowicz and Stephen 2002).

In the next subsections several results will be discussed taking in account different inputs. Each input was normalized using z-score (Jain, Nandakumar and Ross 2005) expressed by the following formula:

$$z_i = \frac{x_i - \mu}{\delta} \quad (7.3)$$

where  $z_i$  is the new input computed,  $x_i$  is the old input and  $\mu$  and  $\delta$  are respectively the mean and the standard deviation of the dataset considered.

Four pairwise classification cases are considered:

- Org&Fun vs Con
- Org vs Con
- Fun vs Con
- Org vs Fun

Each dataset, comprising the two classes considered in each classification case, was divided into five folds and k-fold cross validation ( $k=5$ ) was performed to generalize the results. In 5-fold cross validation data every time a different set for training, validation and test is considered, giving a good estimation of how the classifier generalise the results. Ten runs of the experiment were completed for statistical significance.

We used SVM (Durgesh and Lekha 2010), ANN (Maind and Wankar 2014) and CGP (described in Chapter 2) to generate classifiers distinguishing among the different classes. The ANN used is a feed-forward back propagation network with 20 hidden layers and a tan-Sigmoid transfer function. The geometry of the programs in the population of CGP (chromosomes) has a number of internal nodes equal to three times the number of inputs, with a function set of eleven mathematical functions (+, -, \*, /, mean, min, max, division rest, sin, cos, tan), different number of inputs (described in the following subsections) and one output. Each output is considered a positive response if the value is greater than a certain threshold, negative otherwise; 18 thresholds are considered here, in multiples of 0.1 (0.2, 0.3, 1.80). We chose to use 18 thresholds to facilitate the calculation of AUROC. For each threshold the true positive rate and false positive rate is computed, then the ROC curve is depicted using these two computations for each threshold. The positive class is the first in the classification cases; for example, for the case “Org vs Fun”, the positive class is Org. At each generation the fittest chromosome is selected and the next generation is formed with its mutated versions (mutation rate=0.07). Evolution is stopped when 50000 iterations are reached or when the over fitting occurs. In the next subsections all the inputs considered with the respective classification results will be discussed. The Area Under ROC Curve AUROC (Hanley and McNeil 1982) is used to evaluate the classifiers as the reach and grasp case described in Chapter 5. Also in this case the results are reported in

the form: mean  $\pm$  standard deviation, where the mean and the standard deviation of the AUROCs are computed across the ten runs of the experiment. Both values of train and test set are reported. Then all the results described in the next subsections differ only for the inputs considered. As the test set results are most useful in evaluating the potential of the classifier to generalise the unseen data, from this point forward discussion regarding the results will refer solely to the test set.

### 7.4.1 Maximum for each region

In this subsection the inputs used are the spectrogram maxima extracted for each regions identified in the grid (figure 51). The regions represented in the grid are fourteen and for each region the maximum is extracted corresponding to fourteen inputs. In the following four tables (table 10-13) the results for each classification case are reported.

#### Org&Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
FDS	0.68 $\pm 0.02$	0.75 $\pm$ 0.01	0.65 $\pm 0.02$	0.72 $\pm 0.03$	0.65 $\pm 0.03$	0.71 $\pm 0.04$
EDC	0.60 $\pm 0.02$	0.67 $\pm$ 0.01	0.59 $\pm 0.03$	0.63 $\pm 0.04$	0.57 $\pm 0.01$	0.74 $\pm 0.04$

Table 10. Classification results considering the case Org&Fun vs Con (inputs MAX). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

#### Org vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
FDS	0.69 $\pm 0.02$	0.77 $\pm 0.002$	0.69 $\pm 0.04$	0.72 $\pm 0.03$	0.66 $\pm 0.02$	0.72 $\pm 0.04$
EDC	0.62 $\pm 0.02$	0.68 $\pm$ 0.01	0.63 $\pm 0.02$	0.65 $\pm 0.03$	0.59 $\pm 0.02$	0.72 $\pm 0.05$

Table 11. Classification results considering the case Org vs Con (inputs MAX). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

## Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.60 ± 0.06	0.75 ± 0.01	0.57 ± 0.09	0.68 ± 0.08	0.50 ± 0.07	0.73 ± 0.07
<b>EDC</b>	0.66 ± 0.03	0.77 ± 0.02	0.58 ± 0.08	0.83 ± 0.06	0.53 ± 0.09	0.80 ± 0.05

**Table 12.**Classification results considering the case Fun vs Con (inputs MAX). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

## Org vs Fun

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.70 ± 0.05	0.80 ± 0.004	0.52 ± 0.09	0.69 ± 0.07	0.56 ± 0.03	0.79 ± 0.06
<b>EDC</b>	0.63 ± 0.04	0.77 ± 0.03	0.55 ± 0.07	0.80 ± 0.07	0.59 ± 0.12	0.90 ± 0.03

**Table 13.**Classification results considering the case Org vs Fun (inputs MAX). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

The results reach a maximum of 0.70. In the cases “Org&Fun vs Con” and “Org vs Con”, classifiers evolved using FDS data have better performances respect those evolved using EDC data. SVM evolved the best classifiers reaching 0.68 and 0.69 in the two cases. CGP evolved classifiers with comparable performance to the ones evolved by SVM and ANN in both two cases “Org&Fun vs Con” and “Org vs Con”. The discrimination between Org and Con seems be easier than that between Org&Fun and Con. There is no surprise considering the fact that in functional dystonia the symptoms are not so clear.

The classification between the classes Fun and Con is worse than the previous cases. Obviously this is a challenging task considering the fact that the knowledge of functional dystonia is limited. In this case EDC inputs induced classifiers with better performance with respect to those induced by FDS data. SVM performs better than the other two methods, maybe because it is less influenced by unbalanced data. The difference among SVM, ANN and CGP is more evident in the case “Org vs Fun” where SVM outperforms definitely ANN and CGP generating classifiers reaching a mean AUROC of 0.70, using FDS data as inputs. In this case neither ANN nor CGP are able to evolve classifiers that can distinguish between the two classes. This is no surprise considering that distinguishing between the two classes of patients is a very challenging task due to the similarities. SVM surprisingly generated very good classifiers outperforming also the previous classification cases. In the two cases containing the functional class a big difference between train and test set (overfitting) and increased standard deviations are

present. The overfitting and the increased standard deviations are probably due to the unbalanced data.

### 7.4.2 Mean for each region

The inputs considered are the means extracted for each region in the grid (figure 51). There are fourteen inputs as in the previous case. Four tables (table 14-17) are reported with the results of each classification case.

#### Org&Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
FDS	0.66 ± 0.02	0.73 ± 0.01	0.63 ± 0.03	0.69 ± 0.03	0.60 ± 0.05	0.74 ± 0.03
EDC	0.56 ± 0.05	0.64 ± 0.01	0.55 ± 0.05	0.61 ± 0.03	0.49 ± 0.04	0.69 ± 0.06

Table 14. Classification results considering the case Org&Fun vs Con (inputs MEAN). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

#### Org vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
FDS	0.69 ± 0.03	0.75 ± 0.01	0.65 ± 0.05	0.70 ± 0.03	0.65 ± 0.03	0.78 ± 0.04
EDC	0.60 ± 0.03	0.66 ± 0.01	0.58 ± 0.03	0.63 ± 0.02	0.50 ± 0.05	0.69 ± 0.05

Table 15. Classification results considering the case Org vs Con (inputs MEAN). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

#### Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
FDS	0.57 ± 0.05	0.70 ± 0.02	0.53 ± 0.09	0.73 ± 0.06	0.50 ± 0.04	0.77 ± 0.10
EDC	0.68 ± 0.04	0.74 ± 0.02	0.64 ± 0.10	0.89 ± 0.05	0.58 ± 0.08	0.79 ± 0.07

Table 16. Classification results considering the case Fun vs Con (inputs MEAN). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

## Org vs Fun

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.59 ± 0.05	0.71 ± 0.02	0.57 ± 0.09	0.72 ± 0.08	0.58 ± 0.03	0.78 ± 0.08
<b>EDC</b>	0.70 ± 0.03	0.76 ± 0.03	0.65 ± 0.10	0.83 ± 0.06	0.65 ± 0.12	0.92 ± 0.05

**Table 17. Classification results considering the case Org vs Fun (inputs MEAN). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.**

In the cases “Org&Fun vs Con” and “Org vs Con” the classifiers induced using FDS data are better than those induced using the EDC data. SVM generated classifiers with the best performance in each case. For “Org&Fun vs Con” case CGP evolved classifiers comparable to ANN using FDS data, while ANN and SVM outperform it using EDC data. In “Org vs Con” case, CGP evolved classifiers with almost the same performance of the ones generated by ANN and comparable to the best generated by SVM using FDS data. As in the previous case, the classifiers induced using EDC data by both SVM and ANN, outperform CGP.

Distinguishing Org from Con is easier than distinguishing Org&Fun from Con, as described in the previous subsection where maxima are used as inputs.

The classification cases containing the functional class are the most challenging also in this case. SVM generates the best classifiers for both “Fun vs Con” and “Org vs Fun”. The classifiers obtained using EDC data as inputs are better than the ones considering FDS data. In “Fun vs Con” case, SVM and ANN outperform CGP considering EDC data, while in “Org vs Fun” case SVM outperforms both CGP and ANN which generates comparable classifiers. The overfitting and the increased standard deviation are probably due, as previously, to the unbalanced data.

### 7.4.3 Standard deviation for each region

Again fourteen inputs are used in this case representing the standard deviation extracted for each region of the grid (figure 51). The four tables (table 18-21) with the results are reported below.

### Org&Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.63 ± 0.01	0.73 ± 0.004	0.61 ± 0.04	0.69 ± 0.03	0.62 ± 0.02	0.72 ± 0.03
<b>EDC</b>	0.58 ± 0.04	0.72 ± 0.004	0.61 ± 0.04	0.66 ± 0.02	0.62 ± 0.03	0.72 ± 0.03

**Table 18.**Classification results considering the case Org&Fun vs Con (inputs STD). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

### Org vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.61 ± 0.03	0.74 ± 0.01	0.61 ± 0.04	0.67 ± 0.02	0.62 ± 0.03	0.74 ± 0.03
<b>EDC</b>	0.63 ± 0.02	0.74 ± 0.003	0.65 ± 0.02	0.68 ± 0.004	0.64 ± 0.04	0.74 ± 0.04

**Table 19.**Classification results considering the case Org vs Con (inputs STD). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

### Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.61 ± 0.06	0.63 ± 0.01	0.52 ± 0.11	0.76 ± 0.06	0.59 ± 0.06	0.77 ± 0.06
<b>EDC</b>	0.71 ± 0.04	0.76 ± 0.01	0.54 ± 0.08	0.83 ± 0.06	0.52 ± 0.07	0.75 ± 0.06

**Table 20.**Classification results considering the case Fun vs Con (inputs STD). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

### Org vs Fun

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.56 ± 0.03	0.64 ± 0.02	0.54 ± 0.07	0.74 ± 0.07	0.50 ± 0.05	0.78 ± 0.07
<b>EDC</b>	0.72 ± 0.07	0.78 ± 0.01	0.52 ± 0.06	0.78 ± 0.05	0.55 ± 0.09	0.69 ± 0.10

**Table 21.**Classification results considering the case Org vs Fun (inputs STD). AUROCs mean and standard deviation across the ten runs for FDS and EDC data are reported.

In this case the difference in distinguishing between Org and Con and between Org&Fun and Con is more pronounced. In “Org&Fun vs Con” case the classifiers induced using FDS and EDC data are very similar, reaching a mean AUROC equal respectively to 0.63 and to 0.62. Using EDC data and FDS data, CGP and SVM evolved respectively the best classifiers, comparable to the ones evolved by the other approaches. In “Org vs Con” case the classifiers induced using EDC data are better than the ones evolved using FDS data. All the approaches evolved classifiers comparable, using both FDS and EDC data, with a maximum mean AUROC equal to 0.65.

In both “Fun vs Con” and “Org vs Fun” cases SVM generates the best classifiers reaching a mean AUROC of 0.71 and 0.72 respectively for the two cases and outperforming definitely the other two methods using EDC data. This larger difference can be due to the unbalanced data but it is difficult to explain. In fact neither ANN nor CGP were able to evolve classifiers able to distinguish between the classes, in both cases, either considering FDS data either considering EDC data. As in the previous case, the presence of the overfitting and the increased standard deviation could be due to the unbalanced data.

#### 7.4.4 Max, mean and standard deviation for each region

In this section the number of inputs is greater than the previous cases. The inputs are forty-two, because for each region in the grid (figure 51) the max, the mean and the standard deviation are computed. The four tables (table 22-25) with the results are reported below.

#### Org&Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.61 ± 0.03	0.87 ± 0.01	0.65 ± 0.04	0.76 ± 0.03	0.62 ± 0.04	0.74 ± 0.04
<b>EDC</b>	0.58 ± 0.03	0.77 ± 0.01	0.63 ± 0.03	0.69 ± 0.04	0.59 ± 0.04	0.71 ± 0.04

**Table 22. Classification results considering the case Org&Fun vs Con (inputs MAX, MEAN, STD).**

**AUROC's mean and standard deviation across the ten runs for FDS and EDC data are reported.**

## Org vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.64 ± 0.04	0.82 ± 0.005	0.69 ± 0.03	0.74 ± 0.03	0.65 ± 0.05	0.78 ± 0.03
<b>EDC</b>	0.59 ± 0.04	0.78 ± 0.01	0.65 ± 0.03	0.70 ± 0.02	0.58 ± 0.05	0.75 ± 0.05

**Table 23.**Classification results considering the case Org vs Con (inputs MAX, MEAN, STD).

AUROC's mean and standard deviation across the ten runs for FDS and EDC data are reported.

## Fun vs Con

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.53 ± 0.05	0.83 ± 0.01	0.50 ± 0.07	0.72 ± 0.06	0.54 ± 0.07	0.84 ± 0.06
<b>EDC</b>	0.70 ± 0.07	0.87 ± 0.01	0.57 ± 0.06	0.87 ± 0.06	0.43 ± 0.09	0.75 ± 0.07

**Table 24.**Classification results considering the case Fun vs Con (inputs MAX, MEAN, STD).

AUROC's mean and standard deviation across the ten runs for FDS and EDC data are reported.

## Org vs Fun

DATA	SVM		ANN		CGP	
	Test	Train	Test	Train	Test	Train
<b>FDS</b>	0.70 ± 0.05	0.84 ± 0.01	0.51 ± 0.08	0.75 ± 0.08	0.40 ± 0.07	0.74 ± 0.11
<b>EDC</b>	0.68 ± 0.06	0.86 ± 0.02	0.54 ± 0.11	0.83 ± 0.07	0.59 ± 0.10	0.88 ± 0.06

**Table 25.**Classification results considering the case Org vs Fun (inputs MAX, MEAN, STD).

AUROC's mean and standard deviation across the ten runs for FDS and EDC data are reported.

In this case we tried to improve the results adding information. For both “Org&Fun vs Con” and “Org vs Con” cases the classifiers generated using FDS data have better performances than those generated using EDC data. All the classifiers obtained by the three approaches are comparable, reaching the maximum of mean AUROC equal to 0.69 (ANN using FDS data). The classifiers evolved for “Org vs Con” case have better performances than the ones evolved for “Org&Fun vs Con” case.

In “Fun vs Con” case, using EDC data, SVM outperforms the other approaches that are unable to evolve classifiers to distinguish between the classes. This happens also for classifiers induced by both FDS and EDC data in “Org vs Fun” case. Also

in this case is difficult to explain the large difference in performance between classifiers generated by SVM and the other approaches. Again this could be due to unbalanced data.

### 7.4.5 Fourier Transform comparison

The results using wavelet, as explained in the previous sections, are not satisfactory. For this reason we decided to compute also the Fast Fourier Transform (FFT) of the signal and use this instead that the wavelet transform to classify the different dystonia patients. As features considered for the classification we extracted the mean and the standard deviation plus the first 10 or 20 maxima. The classification is done using only SVM and ANN for the cases “Org&Fun vs Con” and “Org vs Con”. The results are reported in table 26,27,28 and 29. In the following tables the results are in the form: mean  $\pm$  standard deviation, where the mean and the standard deviation of the AUROCs are computed across the ten runs of the experiment. Both values of train and test set are reported. Also using FFT the results do not improve and this is clear examining the following tables. Looking at these results we decided to do not try the classification cases “Org vs Fun” and “Fun vs Con” because an improvement looked improbable given the unbalanced data. The CGP is not used to evolve the classifiers because the results found with SVM and ANN gave a clear indication about the impossibility to classify the different classes. For the oscillatory nature of the EMG signal the use of wavelet is justified to identify both the change in time and in frequency.

#### FDS Org&Fun vs Con

DATA	SVM		ANN	
	Test	Train	Test	Train
<b>10 max + mean + std</b>	0.49 $\pm$ 0.02	0.53 $\pm$ 0.01	0.53 $\pm$ 0.03	0.55 $\pm$ 0.03
<b>10 max + mean + std</b>	0.52 $\pm$ 0.02	0.55 $\pm$ 0.01	0.53 $\pm$ 0.04	0.55 $\pm$ 0.04

Table 26. Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for FDS are reported.

#### FDS Org vs Con

DATA	SVM		ANN	
	Test	Train	Test	Train
<b>10 max + mean + std</b>	0.49 $\pm$ 0.01	0.50 $\pm$ 0.01	0.48 $\pm$ 0.02	0.48 $\pm$ 0.02
<b>10 max + mean + std</b>	0.50 $\pm$ 0.02	0.52 $\pm$ 0.01	0.46 $\pm$ 0.04	0.47 $\pm$ 0.02

Table 27. Classification results considering the case Org vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for FDS are reported.

### EDC Org&Fun vs Con

DATA	SVM		ANN	
	Test	Train	Test	Train
<b>10 max + mean + std</b>	0.47 ± 0.02	0.54 ± 0.01	0.48 ± 0.04	0.54 ± 0.02
<b>10 max + mean + std</b>	0.46 ± 0.03	0.54 ± 0.01	0.50 ± 0.05	0.53 ± 0.03

**Table 28.** Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for EDC are reported.

### EDC Org vs Con

DATA	SVM		ANN	
	Test	Train	Test	Train
<b>10 max + mean + std</b>	0.44 ± 0.05	0.52 ± 0.01	0.46 ± 0.04	0.53 ± 0.02
<b>10 max + mean + std</b>	0.43 ± 0.04	0.52 ± 0.01	0.47 ± 0.04	0.52 ± 0.02

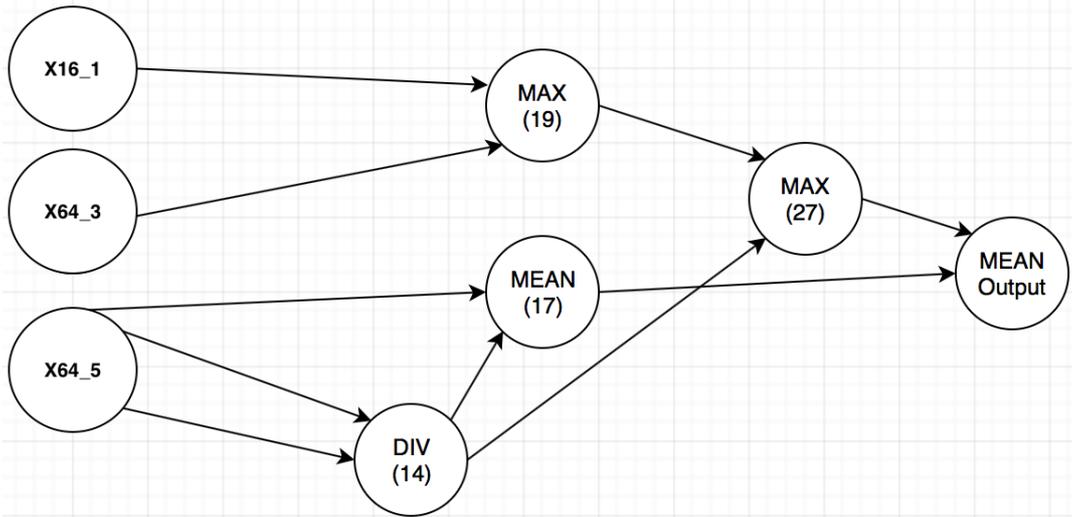
**Table 29.** Classification results considering the case Org&Fun vs Con and using FFT (inputs 10 maxima + mean + std, 20 maxima +mean+ std). AUROCs mean and standard deviation across the ten runs for EDC are reported.

## 7.4.6 CGP classifier Mathematical expressions and Network Diagrams

In this section, as in section 5.3.3, the diagram of an evolved classifier with the derived mathematical expression is reported. Figure 52 shows the best classifier evolved across the cross-validation folds and the generations, for the case “Org vs Con” considering as inputs the 14 maxima extracted from the grid derived by FDS scalogram (section 7.4.1, case “Org vs Con”). The classifier represented has an AUROC equal to 0.80 for the test set. In the diagram only the active nodes are reported to make the graph clearer. Only three of the fourteen inputs have a contribution in the output computation. The order of the inputs is reported in figure 51 showing the grid from which the maxima are extracted. In this case the grid is relative to the FDS scalogram and the inputs used are the maximum extracted from: the first region defined for 16 Hz (x16\_1) and the third and fifth region defined for 64Hz (x64\_3, x64\_5). The mathematical expression derived from the diagram is expression 7.4. In the expression the number of the input was replaced with the short name of the input defined before. The mathematical expression is clear and very easy to implement.

The study of the inputs used, as described in section 5.3.3, can help to understand what the most useful features for each classification case are. Unfortunately identifying the most useful features for each classification case was not possible. In

fact there was not a common set of inputs used by the best classifier for each case across the folds of cross-validation (considering all possible inputs). For this motivation it is not possible to make any conclusion about the most important features for each classification case. The classifier evolved is less complex than that reported for the reach and grasp experiment (section 5.3.3). This is due to the small number of inputs considered (14 in this case, 25 in the reach and grasp case) and consequently to the smaller number of internal nodes considered (equivalent to three times the number of inputs).



**Figure 52. Network diagram of the best classifier evolved for the case Org vs Con considering as inputs the 14 maxima extracted by the FDS scalogram grid. Only the active nodes are reported.**

$$out = mean \left\{ \left[ mean \left( x_{64\_5}, \frac{x_{64\_5}}{x_{64\_5}} \right) \right], \left[ max \left( \frac{x_{64\_5}}{x_{64\_5}}, max(x_{64\_3}, x_{16\_1}) \right) \right] \right\} \quad (7.4)$$

## 7.5 CONCLUSIONS AND FUTURE WORK

In the method proposed, the differences in wavelet spectrogram are used to distinguish among the different classes. Unfortunately, the results evidenced that none of the approaches considered evolved classifiers able to distinguish among the classes. The classes do not seem to have statistically significant differences.

In the first classification case, we tried to distinguish the patients from the healthy subjects (“Org&Fun vs Con”). In this case we achieved the best results considering fourteen inputs, representing the maximum of each region (section 7.4.1). The best results are achieved using FDS data, with evolved classifiers reaching 0.68 for SVM and 0.65 for ANN and CGP.

The second classification case involved the distinction between the organic patients and the healthy subjects (“Org vs Con”). Removing the functional patients from the

patients set brought an improvement in the results for all the input cases considered. This fact is expected because as explained in section 4.4, functional dystonia is not very easy to diagnose and the symptoms might not to be very clear. All the considerations of the previous case (Org&Fun vs Con) are valid for this case: the best classifiers are evolved when the maximum of each region of the grid (14 inputs) is considered as inputs and FDS data are used. In this case, SVM and ANN induced classifiers reaching 0.69, followed by those evolved by CGP, reaching 0.66. There is a slightly improvement with respect to the previous case. Considering the fourteen inputs represented by the mean computed for each region the results are very similar to the ones found considering the maxima. SVM evolved classifiers reaching 0.69 as before, while the ones evolved by ANN and CGP are slightly worse reaching 0.65.

When we consider the functional patients class alone, in the two classification cases “Fun vs Con” and “Org vs Fun”, the results are worse. We can notice that surprisingly there is a big difference between SVM results and the other approaches results in almost all the input cases considered. In almost all the input cases considered the classifiers evolved using EDC data are better than those evolved using FDS results, except for the classification case “Org vs Fun” considering the maximum computed for each region.

The larger difference among the classifiers evolved by SVM and those evolved by other approaches in both cases “Fun vs Con” and “Org vs Fun” considering almost all the input cases, could be due to the fact that the class are unbalanced and maybe in this case SVM tolerates more the unbalanced class than ANN and CGP. The standard deviations computed for classifiers evolved by CGP, ANN and also by SVM in some cases are very high indicating a large variability in the results. The differences between the value of train and test set evidence overfitting, probably due to the unbalanced data. We achieved the best results for both cases “Fun vs Con” and “Org vs Fun” considering as input, the standard deviations computed for each region of the grid derived from EDC data. In “Fun vs Con” SVM induced the best classifiers reaching 0.71, followed by the ones evolved by ANN (0.54) and CGP (0.52). The standard deviations are very high for CGP and ANN evidencing overfitting. In “Org vs Fun” SVM evolved classifiers reaching 0.72, while the ones evolved by CGP and ANN reached 0.59 and 0.52 respectively. In this case all the standard deviations are very high so the results have a high variability.

The results relative to the functional patients class considered alone (not in combination with organic ones) seem to be very variable, so there is less confidence in these results. Probably the unbalanced data affected the classification. Interestingly SVM outperforms the other approaches indicating less sensibility to the unbalanced data. In future it is suggested to repeat the experiment measuring more functional patients in the way that a significant sample is considered. Therefore the recruitment of functional patients could be very difficult considering the symptoms not being very clear and the diagnosis not very easy (section 4.4).

In conclusion, unfortunately as described before, none of the results considered are so high to evidence a statistically significant difference among the classes considered.

These relatively poor results could have a number of reasons:

The problem can depend on the recordings. Section 7.2 described the three problems of using surface EMG: the skin is a possible source of interference and the more body fat the subject has the more EMG data are weak; recordings can be done only on superficial muscles and it is difficult to distinguish the activity among adjacent muscles. These three limitations could affect the recordings and then the results in conjunction with other possible interferences in the room (i.e. light).

The problem could depend on the way in which the recordings were done. The EMG was recorded continuously for all the tasks considered and, as explained in section 7.3.3, the tasks were divided using the shimmers clock. If the clock was not precise the division of the tasks could be wrong. Anyway we are confident enough on the clock precision.

The problem could depend on the task considered; hand opening-closing may not be the best task to address our classification problems. We have chosen this task because it was the task where the muscles considered were more active and it was reasonably easy dividing the single hand-opening repetition.

The muscles considered may not have been affected in a severe way by dystonia and then the recordings among the classes could have been very similar, leading to an increased difficulty in discriminating among the classes.

Another possible explanation could be that the methodology is wrong. To investigate about this hypothesis we used deep learning for image classification, using as inputs the CWT scalogram images relative to each classification case. The network used is convolutional neural networks (CNN) (LeCun, Bengio and Hinton 2015) and the results are computed using the following Matlab toolboxes: Neural Network Toolbox™ and Statistics and Machine Learning Toolbox™. CNN is a deep neural network and then the main problem is the large amount of time needed for the training phase. In order to solve this problem we started from a pre-trained network called AlexNet model. This model is trained on a subset of the ImageNet database (ImageNet. <http://www.image-net.org>), which is used in ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al. 2015). AlexNet has 8 layers with learnable weights: 5 convolutional layers, and 3 fully connected layers. The model is trained on more than a million images and can classify images into 1000 object categories. As a result, the model has learned rich feature representations for a wide range of images. This network was trained again using the EMG spectrogram wavelets images for each of the classification case considered. For simplicity in this case the training and the test set were chosen randomly: 30% of the data were used as test set and the remaining 70% as training sets. In table 30 the results are reported as mean accuracy on the test set for each classification case and for both muscles considered. As in the previous results, ten runs of the experiment were completed and the results are report in the form mean

accuracy±standard deviation, where the mean accuracy and the standard deviation are computed across the ten runs.

<b>Cases</b>	<b>FDS</b>	<b>EDC</b>
<b>Org&amp;Fun vs Con</b>	0.62±0.04	0.54±0.03
<b>Org vs Con</b>	0.64±0.04	0.54±0.02
<b>Fun vs Con</b>	0.49±0.03	0.46±0.08
<b>Org vs Fun</b>	0.54±0.12	0.48±0.08

**Table 30. Deep learning test set results.**  
**The mean accuracy and standard deviation across the ten runs for each classification case and input are reported.**

The problems considering the functional patients as a single class are evident. The results achieved, considering image recognition using CNN, are not better than the results achieved with the methodology described in this work. This fact seems to highlight that there are no significant difference in the data recorded, supporting the view that our results are affected by the data itself instead by the methodology used.

In future work we advise to repeat the experiment separating each single task and including patient with hand dystonia only, in order to be sure that the disease seriously affects the muscles recorded. Also a change of the task can be considered recording other relevant muscles. The use of intramuscular EMG might be useful to increase the measurements precision.

Future studies, following our suggestions, will overcome the limitations of this study with a consequent improvement of the results.

# CHAPTER 8: CONCLUSIONS AND FUTURE WORK

## 8.1 RESEARCH SUMMARY

The principal aim of this work was to improve the diagnosis and the monitoring of two neurological disorders: Parkinson's disease and dystonia. For both of them no objective clinical assessments currently exist and the diagnosis is based on subject clinical assessment that is often not very reliable. Actually the misdiagnosis rate of Parkinson's disease is about 25% (Playfer 1997) that is equivalent to saying that 25% of people diagnosed with Parkinson's disease are ultimately found to have another similar condition. This misdiagnosis rate is high, due to the similarity of the symptoms among Parkinson's disease and other neurological conditions. Dystonia (or more accurately organic dystonia) is actually diagnosed evaluating its core motor features and temporal evolution (Albanese et al. 2013) with a current misdiagnosis rate of 25-52% (Pal 2011). The misdiagnosis rate is due to the existence of several atypical forms of organic dystonia that are not easy to separate from functional dystonia. The difference between organic and functional dystonia is explained in details in section 4.4. The distinction between the two forms is crucial to manage the conditions with the best drugs. An objective way to diagnose Parkinson's disease and dystonia is necessary, in order to reduce the respectively misdiagnosis rates.

The identification of new features characterizing Parkinson's disease and dystonia could help in improving the diagnosis of these conditions, reducing the misdiagnosis rate. Twenty-five features characterizing Parkinson's disease patients and healthy subjects were computed from the kinematic reach and grasp data (section 5.2.5). One of the novelties of this study is the computation of this set of features by merging two set of features considered respectively in Alberts' and Caselli's studies (Alberts et al. 2000, Caselli et al. 1999). The features were used as input to the CGP in order to evolve classifiers able to distinguish not only between Parkinson's disease patients and healthy subjects, but also among Parkinson's disease patients affected by two different kinds of cognitive impairments. The cognitive decline of Parkinson's disease is then assessed with an artificial intelligence technique. The assessment of cognitive decline is important in monitoring the progression of the disease modifying drugs if necessary and also to identify the mild cognitive impairment (section 4.4), which treated it in the proper way, can prevent dementia.

The EMG data of dystonia patients and healthy subjects collected in the hand opening-closing experiment were pre-processed with Morse wavelet (section 6.1.3) and divided in a grid (figure 51), then different features were extracted (sections 7.4.1-7.4.4). The pre-processing of the EMG data with Morse wavelet together with the extraction of features derived represent two novelties of this study.

The features extracted from both reach and grasp and hand opening-closing experiments were used as inputs to the evolutionary algorithms to evolve the best classifiers. The performance of the classifiers evolved by the evolutionary algorithms, in this case the Cartesian Genetic Programming (section 2.4) are compared to others obtained using two well-known machine learning techniques: Support Vector Machine (Durgesh and Lekha 2010) and Artificial Neural Networks (Maind and Wankar 2014). The main objective was to show the potential of artificial intelligence techniques to improve the monitoring and the diagnosis of the two neurological disorders. Evolutionary algorithms in particular are useful to give a better insight about the features characterising the cognitive decline in Parkinson's disease and the differentiation between organic and functional dystonia. In fact these algorithms, evolve classifiers that do not necessarily use all the inputs considered but only the most useful, selecting the most discriminative ones (section 5.3.3 and section 7.4.5). The use of artificial intelligence techniques could help also in reducing the costs associated to the neurological diseases that, as explained in section 1.2, can be very high. Both of the experiments made in fact are easy to do, non-invasive, cost efficient and could design a more objective easy test for the diagnosis.

The two main aims, as described in section 1.4, were:

1. Finding a classifier able to distinguish between Parkinson's disease patients and healthy subjects, but also among different subgroups of patients that present different cognitive impairments.
2. Finding a classifier able to distinguish between dystonia patients and healthy subjects but also between two different subgroups of patients (organic and functional ones).

In the next section all the results will be examined to verify if this study fulfilled its aims. The research question will be revisited in section 8.2.3.

## **8.2 CONCLUSIONS AND FUTURE WORK**

The two aims of the research, described in section 8.1, denoted the division of the study in two different parts: the part relative to Parkinson's disease patients and that pertaining to the dystonia patients. In the first part a reach and grasp experiment (Chapter 5) is considered while in the second part a hand opening-closing experiment (Chapter 7) is considered. In this section we give the final considerations with suggestions for future work for both parts (respectively section 8.2.1-8.2.20), considering the results. Finally, in section 8.2.3, all the conclusions are summarized, revisiting the research question (section 1.6).

## 8.2.1 Parkinson's disease classification

In reach and grasp experiment different classifiers were evolved to distinguish among different Parkinson's disease patients and healthy subjects. The classes considered are: Parkinson's disease patients with normal cognition (PD-NC), Parkinson's disease patients with mild cognitive impairment (PD-MCI), Parkinson's disease patients with dementia (PDD), all Parkinson's disease patients together (PD-NC, PD-MCI and PDD) and healthy age-matched subjects (Controls). The best classifiers are evolved for the case "PDD vs Controls" while the worse for the case "PD-NC versus PD-MCI". These results were expected because: dementia can badly affect the movements evidencing the differences between patients and controls; while mild cognitive impairment instead does not affect badly the patient's life therefore the difference between normal cognition patients and those affected by mild cognitive impairment is not pronounced. More pronounced were the differentiations in the cases "PD-NC vs PDD" and "PD-MCI vs PDD", this is expected as well because as described before, dementia can badly affect the movements, and therefore the differences among the classes are much evident. It is important also to remember that the data considered are relative to the experiment done with eyes closed, and then the subjects are guided by memory which is severely affected by dementia.

Cartesian Genetic Programming evolved classifiers comparable to those obtained by the other approaches in the cases including controls while the other approaches outperformed CGP in the comparisons among the patient subgroups, especially in "PD-NC vs PD-MCI". The results suggest that CGP is not the best approach, but its big advantage is that it can translate classifiers into simple mathematical expressions and make their implementation very easy. In section 5.3.3 a diagram of an evolved classifier with its derived mathematical expression is reported. The graphical and mathematical representations of the classifier are used to understand how the data are used and can give a better insight about the nature of the data.

This study represents a start point for the diagnosis and the monitoring of cognitive decline in Parkinson's disease. The features computed evidenced some differences among the classes considered, with the best results achieved for the classification cases containing patients affected by dementia. Patients affected by dementia seem have the most difficulties in performing the "reach and grasp" experiment with closed eyes.

This study has however two main limitations:

1. The repetitions of the experiment for each subject were treated as different samples. This could affect the study because the same subject can perform the experiment in a similar way, reducing the variability. If we used the mean across the repetitions, the data were not enough to evolve classifiers. This choice was justified observing a great variability across repetitions for the patients.

2. The patients were on medication to limit the motor symptoms of Parkinson's disease, obviously the medications could influence the results improving patients' performances.

In order to overcome the two limitations it is possible to repeat the reach and grasp experiment, with closed eyes, including more subjects to increase the variability and also patients that are not on medications to evaluate their effects on the experiment.

Another point of discussion is the use of the grasp data recorded with the glove sensors. In this study the grasp data are not used for the motivations explained in section 5.2.6. Therefore the use of the grasp data could be very useful to identify new features useful for the classification. For future work a new reach and grasp experiment could be considered using electromagnetic sensors, as that placed on the wrist (figure 23), placed on the fingers. In this way it is possible to record the position of the fingers during the experiment permitting the computation of the distance between fingers and then some useful features such as maximum aperture, time of maximum aperture, time of opening onset can be obtained.

The features chosen are obtained merging feature sets found in two previous studies (Caselli et al. 1999, Alberts et al. 2000). These 25 features are considered enough to represent the reach and grasp and then differentiate among the classes. Moreover summarizing the data available with a limited set of features can cause a loss of useful information very important for the classification. Further investigations are needed to find a way to use the raw data instead of only a limited set of features. In conclusion this study is a starting point for the classification of cognitive decline in Parkinson's disease; some differences among the different subgroups of patients are evidenced, highlighting overall the difficulties of the patients affected by dementia.

### **8.2.2 Dystonia classification**

In the hand opening-closing experiment, EMG data were pre-processed with Morse wavelet (section 6.1.3), divided in a grid (figure 51) and then different features were extracted. The class considered are: healthy subjects (Con), organic dystonia patients (Org), functional dystonia patients (Fun) and all the dystonia patients (Org&Fun). Using the grid (figure 51) we extracted different features for each region such as mean, maximum and standard deviation. Unfortunately in none of the cases considered statistical significance differences between the classes are found. In classification cases, which contain functional patients ("Fun vs Con" and "Org vs Fun"), SVM outperforms the other approaches that are unable to differentiate the classes. This distance between SVM and the other approaches along with the high standard deviations, make us unsure about the results. The unbalanced datasets could explain this high variability; in fact the functional subjects recorded are very few with respect to the other classes.

There are several factors that could affect the results. The methodology could be wrong or the problem could be in the data.

In order to verify if the problem was the methodology we considered the scalogram images and used the deep learning Matlab toolboxes for images classification. In table 30 the test set accuracies relative to the convolutional neural network are reported. The results are no better than those found with our methodology, highlighting the difficulty to distinguish among functional patients and other classes.

If the problem is not the methodology then probably it is in the data that do not evidence the differences between classes. There are a lot of factors that could influence the quality of the EMG signal. The surface EMG, as described in section 7.2, are not invasive as the intramuscular ones but have three main limitations: the skin is a possible source of interference and more body fat a subject has the weak the EMG signal are; recordings can be done only on superficial muscles and it is difficult to distinguish the activity among adjacent muscles. The interferences due to the skin and the adjacent muscles could affect the quality of the signal. The position of the electrodes on the muscles is very important to minimize the adjacent muscles interferences. In order to record a good quality signal the electrode has to be placed at the belly of the muscle (longitudinal midline) (Delsys technical note 101). This position is very difficult to determine because it is not possible to see the exact position of the muscles under the skin. The inexperience of the doctor that recorded the EMG data at Monash Medical Centre (MMC) in Melbourne (Australia) could lead to some errors in positioning the electrodes. Furthermore the recording was blind; this means that was not possible to check the signal during the recording detecting possible errors.

The hand opening-closing movement could be not the best choice to appreciate the differences among the classes considered, but it was chosen because was one of the movements where the muscles considered were more active and also it was easy to divide it in the single hand-opening trials using the kinematic data. Also the muscles considered could not be badly affected by dystonia making the differences not so clear. All these factors could influence our recording and then our results. In order to verify our methodology a new experiment made by an expert person could be considered, using equipment that permits to check the signal during the recording.

For future work, different movements could be considered recording all the specific muscles involved in the movements. Then all the different data for each task could be analysed with the methodology described, in order to verify if there are particular tasks more useful for the classifications. More functional patients should be included in future experiments to improve our classification results solving the problem of unbalanced data.

In order to verify the hand opening-closing movement itself, it could be useful to repeat the experiment including only patients affected by hand dystonia to be sure that the muscles recorded are affected by the disease.

The previous suggestions, hopefully, will be useful in future to overcome the limitations of this study and then to improve the results obtained.

### **8.2.3 Final conclusions**

In section 8.2.1 and 8.2.2 we reported the conclusions and the suggestions for future work for both parts of the study. Using these considerations we can now revisit the research question defined in section 1.6 as:

***“Can Evolutionary algorithms provide a means for monitoring and diagnosing of specific neurological disorders?”***

The main question is: can the results found in this study justify the use of the Evolutionary algorithms as a mean for monitoring and diagnosing of the two specific neurological disorders considered?

In the first part of the study the Evolutionary algorithms are used to diagnose Parkinson’s disease but also to diagnose and to monitor the cognitive decline associated to the disease. The results evidence that CGP evolved classifiers are comparable to the other approaches in differentiating between controls and patients, while the other approaches evolved better classifiers in differentiating among patients affected by different cognitive impairments. Also if the other approaches outperform CGP in some cases, the advantage of CGP is to translate the program in a simple mathematical expression making its implementation very easy and cost efficient. Another potential advantage of CGP or in general of the evolutionary algorithms, is the fact that they do not use necessarily all the inputs and can give indications about the most important features. Details about the classifiers diagram and their derived mathematical expression are given in section 5.3.3. Unfortunately during the cross-validation it was not possible to find a common subset of features used by the best classifier for each classification case, then it is not possible to determine the best features. The advantage of CGP could be appreciated better if further investigations will lead to use the raw data instead only a subset of features; in fact CGP could determine distinctive movement patterns in the raw data as demonstrated in previous studies (Smith et al. 2007, Lacy et al. 2013, Lones et al. 2014).

In the second part of the study evolutionary algorithms are applied to find differences in EMG data among organic dystonia patients, functional dystonia patients and healthy subjects. Unfortunately neither the classifiers evolved by CGP, nor those obtained using the other approaches were able to find statistical significance differences among the classes considered. The motivations of the bad results, as explained in section 8.2.2, seemed to depend on the data recorded that do not evidence differences among the classes. In order to verify the quality of the methodology, a new experiment has to be considered ensuring a good quality of the signal. Some suggestions for future work, useful to overcome the limitations of this

study are given in section 8.2.2. Then the judgement of the methodology has to be postponed until a new good quality data set including more functional dystonia patients is available.

The results relative to the Parkinson's disease evidence the potential of the evolutionary algorithms as a mean of diagnosing and monitoring of the disease. This technique can be very useful to help doctors in the diagnosis itself but also in monitoring the cognitive decline associated. In dystonia instead the classifiers evolved are not able to distinguish among the classes. The problem seems to depend on the data as discussed previously. We can conclude that this study shows a potential of using the evolutionary algorithms in the context of diagnosing and monitoring neurological disorders. Further investigations are needed to confirm the results and verify the methodology. However, the results relative to Parkinson's disease are encouraging and can justify the use of this technique.

The hope of the author is that the study can represent an incentive for further investigations on the use of evolutionary algorithms as a means to help doctors in the diagnosing and monitoring of neurological disorders. The use of CGP can be useful to understand better how the data are used in classification (diagrams in section 5.3.3 and 7.4.2) and then have a better insight about the nature of the data. The technique can be applied also to other neurological disorders such as Alzheimer disease or multiple sclerosis investigating the potential of artificial intelligence techniques. Hopefully this work will be useful for future works that will overcome limitations and improve the results.

## REFERENCES

- Addison, P. S. 2002. The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. Institute of Physics Publishing.
- Addison, P. S., J. N. Watson, G. R. Clegg, P. A. Steen & C. E. Robertson (2002) Finding coordinated atrial activity during ventricular fibrillation using wavelet decomposition. *IEEE Engineering in Medicine and Biology Magazine*, 21, 58-65.
- Aereospace MATLAB Toolbox. 2017. <http://mathworks.com/help/aerotbx/index.html>.
- Aguiar - Conraria, L. & M. J. Soares (2014) The continuous wavelet transform: moving beyond uni - and bivariate analysis. *Journal of Economic Surveys*, 28, 344-375.
- Ahmed, F. E. (2005) Artificial neural networks for diagnosis and survival prediction in colon cancer. *Molecular cancer*, 4, 29.
- Aho, A. V. & J. D. Ullman. 1983. *Data structures and algorithms*. Pearson.
- Akansu, A. N., P. Duhamel, X. Lin & M. De Courville (1998) Orthogonal transmultiplexers in communication: A review. *IEEE Transactions on signal processing*, 46, 979-995.
- Akansu, A. N. & M. J. Medley. 2006. *Wavelet, subband and block transforms in communications and multimedia*. Springer Science & Business Media.
- Akansu, A. N. & M. J. T. Smith. 2012. *Subband and wavelet transforms: design and applications*. Springer Science & Business Media.
- Albanese, A., K. Bhatia, S. B. Bressman, M. R. DeLong, S. Fahn, V. S. C. Fung, M. Hallett, J. Jankovic, H. A. Jinnah & C. Klein (2013) Phenomenology and classification of dystonia: a consensus update. *Movement Disorders*, 28, 863-873.
- Alberts, J. L. (1997) The coordination and phasing of a bimanual prehension task: the influence of Parkinson's disease.
- Alberts, J. L., M. Saling, C. H. Adler & G. E. Stelmach (2000) Disruptions in the reach-to-grasp actions of Parkinson's patients. *Experimental Brain Research*, 134, 353-362.
- Aldroubi, A. & M. Unser. 1996. *Wavelets in medicine and biology*. CRC press.
- Amboni, M., P. Barone & J. M. Hausdorff (2013) Cognitive contributions to gait and falls: evidence and implications. *Movement Disorders*, 28, 1520-1533.
- Amboni, M., P. Barone, L. Ippariello, I. Lista, R. Tranfaglia, A. Fasano, M. Picillo, C. Vitale, G. Santangelo & V. Agosti (2012) Gait patterns in Parkinsonian patients with or without mild cognitive impairment. *Movement Disorders*, 27, 1536-1543.
- Apetauerova, D., C. M. Schirmer, J. L. Shils, J. Zani & J. E. Arle (2010) Successful bilateral deep brain stimulation of the globus pallidus internus for persistent status dystonicus and generalized chorea. *Journal of Neurosurgery*, 113, 634-638.
- Balint, B. & K. P. Bhatia (2014) Dystonia: an update on phenomenology, classification, pathogenesis and treatment. *Current Opinion in Neurology*, 27.
- Banzhaf, W. 1993. Genetic programming for pedestrians. 628.
- Bartosch - Härlid, A., B. Andersson, U. Aho, J. Nilsson & R. Andersson (2008) Artificial neural networks in pancreatic disease. *British Journal of Surgery: Incorporating European Journal of Surgery and Swiss Surgery*, 95, 817-826.
- Batista, G. E., R. C. Prati & M. C. Monard (2004) A study of the behavior of several methods for balancing machine learning training data. *ACM Sigkdd Explorations Newsletter*, 6, 20-29.
- Bentivoglio, A. R., M. Loi, E. M. Valente, T. Ialongo, P. Tonali & A. Albanese (2002) Phenotypic variability of DYT1 - PTD: Does the clinical spectrum include psychogenic dystonia? *Movement disorders*, 17, 1058-1063.

- Bermingham, S. L., A. Cohen, J. Hague & M. Parsonage (2010) The cost of somatisation among the working - age population in England for the year 2008–2009. *Mental Health in Family Medicine*, 7, 71.
- Bhowan, U., M. Johnston, M. Zhang & X. Yao (2013) Evolving diverse ensembles using genetic programming for classification with unbalanced data. *IEEE Transactions on Evolutionary Computation*, 17, 368-386.
- Bianchi, L., M. Dorigo, L. M. Gambardella & W. J. Gutjahr (2009) A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing*, 8, 239-287.
- Bracewell, R. (2000) Heaviside's Unit Step Function. *The Fourier Transform and Its Applications*, 61-65.
- Bradley, A. P. (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30, 1145-1159.
- Brain Foundation. 2018. <http://brainfoundation.org.au/disorders/dystonia>.
- Brameier, M. & W. Banzhaf (2001) A comparison of linear genetic programming and neural networks in medical data mining. *IEEE Transactions on Evolutionary Computation*, 5, 17-26.
- Brameier, M. F. & W. Banzhaf. 2007. *Linear genetic programming*. Springer Science & Business Media.
- Bressman, S. B., D. De Leon, M. F. Brin, N. Risch, R. E. Burke, P. E. Greene, H. Shale & S. Fahn (1989) Idiopathic dystonia among ashkenazi jews: Evidence for autosomal dominant inheritance. *Annals of Neurology*, 26, 612-620.
- Broeders, M., R. M. A. De Bie, D. C. Velseboer, J. D. Speelman, D. Muslimovic & B. Schmand (2013) Evolution of mild cognitive impairment in Parkinson disease. *Neurology*, 81, 346-352.
- Broughton, S. A. (1998) Wavelet based methods in image processing. *Rose-Hulman Institute of Technology, Winter Quarter*.
- Burke, R. E., S. Fahn, C. D. Marsden, S. B. Bressman, C. Moskowitz & J. Friedman (1985) Validity and reliability of a rating scale for the primary torsion dystonias. *Neurology*, 35, 73-73.
- Busch, P., T. Heinonen & P. Lahti (2007) Heisenberg's uncertainty principle. *Physics Reports*, 452, 155-176.
- Carmona, R., W.-L. Hwang & B. Torresani. 1998. *Practical Time-Frequency Analysis: Gabor and wavelet transforms, with an implementation in S*. Academic Press.
- Carmona, R. A., W. L. Hwang & B. Torr sani (1997) Characterization of signals by the ridges of their wavelet transforms. *IEEE transactions on signal processing*, 45, 2586-2590.
- (1999) Multiridge detection and time-frequency reconstruction. *IEEE transactions on signal processing*, 47, 480-492.
- Caselli, R. J., G. E. Stelmach, J. N. Caviness, D. Timmann, T. Royer, B. F. Boeve & J. E. Parisi (1999) A kinematic study of progressive apraxia with and without dementia. *Mov Disord*, 14, 276-87.
- Castiello, U., K. Bennett, C. Bonfiglioli, S. Lim & R. F. Peppard (1999) The reach-to-grasp movement in Parkinson's disease: response to a simultaneous perturbation of object position and object size. *Experimental brain research*, 125, 453-462.
- Castiello, U., K. M. B. Bennett & C. Mucignat (1993) The reach to grasp movement of blind subjects. *Experimental Brain Research*, 96, 152-162.
- Castiello, U., G. E. Stelmach & A. N. Lieberman (1993) Temporal dissociation of the prehension pattern in Parkinson's disease. *Neuropsychologia*, 31, 395-402.
- Chartrand, G., L. Lesniak & P. Zhang. 2010. *Graphs & digraphs*. CRC press.
- Clegg, J., J. A. Walker & J. F. Miller. 2007. A New Crossover Technique for Cartesian Genetic Programming.

- Cognitive Impairment Parkinson's Disease Foundation. [http://www.pdf.org/en/cognitive\\_impairment\\_pd](http://www.pdf.org/en/cognitive_impairment_pd).
- Cooper, J. A., H. J. Sagar, P. Tidswell & N. Jordan (1994) Slowed central processing in simple and go/no-go reaction time tasks in Parkinson's disease. *Brain*, 117 ( Pt 3), 517-29.
- Cosgrove, J. (2016) Investigating reach and grasp in Parkinson's disease cognitive impairment. Cosgrove, J., C. Picardi, S. L. Smith, S. Jamieson & J. E. Alty. 2016. How does cognition affect reaching in Parkinson's disease? , S459-S459. WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA.
- Cramer, N. L. A representation for the adaptive generation of simple sequential programs. 183-187.
- Criswell, S., C. Sterling, L. Swisher, B. Evanoff & B. A. Racette (2010) Sensitivity and specificity of the finger tapping task for the detection of psychogenic movement disorders. *Parkinsonism & related disorders*, 16, 197-201.
- Darwin, C. & W. F. Bynum. 2009. *The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life*. AL Burt.
- Daubechies, I. 1992. *Ten lectures on wavelets*. SIAM.
- De Vries, P. M., K. L. Leenders, J. H. Van Der Hoeven, B. M. De Jong, A. J. Kuiper & N. M. Maurits (2007) Abnormal surface EMG during clinically normal wrist movement in cervical dystonia. *European journal of neurology*, 14, 1244-1250.
- Defazio, G., P. Livrea, G. Guanti, V. Lepore & E. Ferrari (1993) Genetic Contribution to Idiopathic Adult-Onset Blepharospasm and Cranial-Cervical Dystonia. *European Neurology*, 33, 345-350.
- Delprat, N., B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian & B. Torresani (1992) Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies. *IEEE transactions on information theory*, 38, 644-664.
- Delsys technical note 101. [https://www.delsys.com/Attachments\\_pdf/TN101%20-%20EMG%20Sensor%20Placement-web.pdf](https://www.delsys.com/Attachments_pdf/TN101%20-%20EMG%20Sensor%20Placement-web.pdf).
- Deo, N. 2017. *Graph theory with applications to engineering and computer science*. Courier Dover Publications.
- Deuschl, G. & R. Elble (2009) Essential tremor—neurodegenerative or nondegenerative disease towards a working definition of ET. *Movement Disorders*, 24, 2033-2041.
- Docherty, M. J. & D. J. Burn (2010) Parkinson's disease dementia. *Current neurology and neuroscience reports*, 10, 292-298.
- Durgesh, K. S. & B. Lekha (2010) Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology*, 12, 1-7.
- Dystonia Medical Research Foundation. 2018. <https://www.dystonia-foundation.org/what-is-dystonia/forms-of-dystonia/secondary-dystonia/more-on-secondary-dystonia>.
- Eigen, M. (1973) Ingo Rechenberg Evolutionsstrategie Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. *mit einem Nachwort von Manfred Eigen*, Friedrich Frommann Verlag, Struttgart-Bad Cannstatt, 45, 46-47.
- Enescu, B., K. Ito & Z. R. Struzik (2006) Wavelet-based multiscale resolution analysis of real and simulated time-series of earthquakes. *Geophysical Journal International*, 164, 63-74.
- Espay, A. J. & A. E. Lang (2015) Phenotype-specific diagnosis of functional (psychogenic) movement disorders. *Current neurology and neuroscience reports*, 15, 32.
- Espay, A. J., F. Morgante, J. Purzner, C. A. Gunraj, A. E. Lang & R. Chen (2006) Cortical and spinal abnormalities in psychogenic dystonia. *Annals of neurology*, 59, 825-834.
- Fahn, S. (2008) The history of dopamine and levodopa in the treatment of Parkinson's disease. *Mov Disord*, 23 Suppl 3, S497-508.

- Fan, R.-E., P.-H. Chen & C.-J. Lin (2005) Working set selection using second order information for training support vector machines. *Journal of machine learning research*, 6, 1889-1918.
- Findley, L. J. (2007) The economic impact of Parkinson's disease. *Parkinsonism Relat Disord*, 13 Suppl, S8-s12.
- Fineberg, N. A., P. M. Haddad, L. Carpenter, B. Gannon, R. Sharpe, A. H. Young, E. Joyce, J. Rowe, D. Wellsted, D. J. Nutt & B. J. Sahakian (2013) The size, burden and cost of disorders of the brain in the UK. *Journal of Psychopharmacology (Oxford, England)*, 27, 761-770.
- Fogel, L. J., A. J. Owens & M. J. Walsh (1966) Artificial intelligence through simulated evolution.
- Freitas, A. A. 2009. A review of evolutionary algorithms for data mining. In *Data Mining and Knowledge Discovery Handbook*, 371-400. Springer.
- Friedberg, R. M. (1958) A learning machine: Part I. *IBM Journal of Research and Development*, 2, 2-13.
- Friedberg, R. M., B. Dunham & J. H. North. 1959. *A learning machine: Part II*.
- Friedman, A. & S. Fahn (1986) Spontaneous remissions in spasmodic torticollis. *Neurology*, 36, 398-398.
- Fukui, T. & T. Inui (2013) Utilization of visual feedback of the hand according to target view availability in the online control of prehension movements. *Human movement science*, 32, 580-595.
- Ganos, C., M. J. Edwards & K. P. Bhatia (2014) The Phenomenology of Functional (Psychogenic) Dystonia. *Movement Disorders Clinical Practice*, 1, 36-44.
- Gary, K. (2004) Electromyographic kinesiology. *Research methods in biomechanics. Human Kinetics Publ, Champaign*.
- Geman, O. 2013. Nonlinear dynamics, artificial neural networks and neuro-fuzzy classifier for automatic assessing of tremor severity. 1-4. IEEE.
- Gentilucci, M. & A. Negrotti (1999) The control of an action in Parkinson's disease. *Experimental Brain Research*, 129, 269-277.
- Geyer, H. L. & S. B. Bressman (2006) The diagnosis of dystonia. *The Lancet Neurology*, 5, 780-790.
- Ghumbre, S., C. Patil & A. Ghatol. Heart disease diagnosis using support vector machine.
- Gibb, W. R. & A. J. Lees (1988) The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease. *J Neurol Neurosurg Psychiatry*, 51, 745-52.
- Giovannoni, G., J. van Schalkwyk, V. U. Fritz & A. J. Lees (1999) Bradykinesia akinesia inco-ordination test (BRAIN TEST): an objective computerised assessment of upper limb motor function. *J Neurol Neurosurg Psychiatry*, 67, 624-9.
- Goetz, C. G. 2011. The History of Parkinson's Disease: Early Clinical Descriptions and Neurological Therapies. In *Cold Spring Harb Perspect Med*.
- Goetz, C. G., B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez - Martin, W. Poewe, C. Sampaio, M. B. Stern & R. Dodel (2008) Movement Disorder Society - sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS - UPDRS): Scale presentation and clinimetric testing results. *Movement disorders*, 23, 2129-2170.
- Gowers, W. R. 1898. *A manual of diseases of the nervous system*. P. Blakiston, Son & Company.
- Grosan, C. & A. Abraham (2008) A New Approach for Solving Nonlinear Equations Systems. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 38, 698-714.
- Gustavsson, A., F. Svensson M Fau - Jacobi, C. Jacobi F Fau - Allgulander, J. Allgulander C Fau - Alonso, E. Alonso J Fau - Beghi, R. Beghi E Fau - Dodel, M. Dodel R Fau - Ekman, C. Ekman M Fau - Faravelli, L. Faravelli C Fau - Fratiglioni, B. Fratiglioni L Fau - Gannon, D. H. Gannon B Fau - Jones, P. Jones Dh Fau - Jennum, A. Jennum P Fau - Jordanova, L. Jordanova A Fau - Jonsson, K. Jonsson L Fau - Karampampa, M. Karampampa K Fau -

- Knapp, G. Knapp M Fau - Kobelt, T. Kobelt G Fau - Kurth, R. Kurth T Fau - Lieb, M. Lieb R Fau - Linde, C. Linde M Fau - Ljungcrantz, A. Ljungcrantz C Fau - Maercker, B. Maercker A Fau - Melin, M. Melin B Fau - Moscarelli, A. Moscarelli M Fau - Musayev, F. Musayev A Fau - Norwood, M. Norwood F Fau - Preisig, M. Preisig M Fau - Pugliatti, J. Pugliatti M Fau - Rehm, L. Rehm J Fau - Salvador-Carulla, B. Salvador-Carulla L Fau - Schlehofer, R. Schlehofer B Fau - Simon, H.-C. Simon R Fau - Steinhausen, L. J. Steinhausen Hc Fau - Stovner, J.-M. Stovner Lj Fau - Vallat, P. Vallat Jm Fau - Van den Bergh, J. Van den Bergh P Fau - van Os, P. van Os J Fau - Vos, W. Vos P Fau - Xu, H.-U. Xu W Fau - Wittchen, B. Wittchen Hu Fau - Jonsson, J. Jonsson B Fau - Olesen & J. Olesen (2011) Cost of disorders of the brain in Europe 2010.
- Haddad, S. A. P. & W. A. Serdijn. 2009. Wavelet versus Fourier Analysis. In *Ultra Low-Power Biomedical Signal Processing: An Analog Wavelet Filter Approach for Pacemakers*, eds. S. A. P. Haddad & W. A. Serdijn, 33-50. Dordrecht: Springer Netherlands.
- Haibo He, Yang Bai, E. A. Garcia & Shutao Li. 2008. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322-1328.
- Hanley, J. A. & B. J. McNeil (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29-36.
- Hassan, A., S. Vallabhajosula, L. B. Zahodne, D. Bowers, M. S. Okun, H. H. Fernandez & C. J. Hass (2014) Correlations of apathy and depression with postural instability in Parkinson disease. *Journal of the neurological sciences*, 338, 162-165.
- He, H. & E. A. Garcia (2009) Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21, 1263-1284.
- Hinson, V. K., E. Cubo, C. L. Comella, C. G. Goetz & S. Leurgans (2005) Rating scale for psychogenic movement disorders: scale development and clinimetric testing. *Movement disorders*, 20, 1592-1597.
- Hoehn Mm Fau - Yahr, M. D. & M. D. Yahr (1967) Parkinsonism: onset, progression and mortality.
- Holland, J. H. 1992. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.
- Hughes, A. J., S. E. Daniel, S. Blankson & A. J. Lees (1993) A clinicopathologic study of 100 cases of Parkinson's disease. *Arch Neurol*, 50, 140-8.
- International Parkinson and Movement Disorder Society (MDS). 2013. <http://www.movementdisorders.org/MDS/News/Online-Web-Edition/Archived-Editions/A-History-of-Dystonia.htm>.
- Iranmanesh, S. & E. Rodriguez-Villegas (2017) A 950 nW Analog-Based Data Reduction Chip for Wearable EEG Systems in Epilepsy. *IEEE Journal of Solid-State Circuits*, 52, 2362-2373.
- Isenberg, C. & B. Conrad (1994) Kinematic properties of slow arm movements in Parkinson's disease. *Journal of Neurology*, 241, 323-330.
- Jain, A., K. Nandakumar & A. Ross (2005) Score normalization in multimodal biometric systems. *Pattern Recognition*, 38, 2270-2285.
- Jakobson, L. S. & M. A. Goodale (1991) Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Experimental Brain Research*, 86, 199-208.
- Jamshidi, M. (2003) Tools for intelligent control: fuzzy controllers, neural networks and genetic algorithms. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 361, 1781-1808.
- Jankovic, J. (2002) Essential tremor: a heterogenous disorder. *Movement disorders*, 17, 638-644.
- (2008) Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry*, 79, 368-76.

- Jankovic, J., K. S. Schwartz & W. Ondo (1999) Re-emergent tremor of Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 67, 646-650.
- Janvin, C. C., J. P. Larsen, D. Aarsland & K. Hugdahl (2006) Subtypes of mild cognitive impairment in Parkinson's disease: progression to dementia. *Movement Disorders*, 21, 1343-1349.
- Japkowicz, N. & S. Stephen (2002) The class imbalance problem: A systematic study. *Intelligent data analysis*, 6, 429-449.
- Jayne, D., A. J. Lees & G. M. Stern (1984) Remission in spasmodic torticollis. *Journal of Neurology, Neurosurgery & Psychiatry*, 47, 1236.
- Jeannerod, M. (1984) The timing of natural prehension movements. *Journal of motor behavior*, 16, 235-254.
- (1986) The formation of finger grip during prehension. A cortically mediated visuomotor pattern. *Behavioural brain research*, 19, 99-116.
- Jin, K., J. W. Simpkins, X. Ji, M. Leis & I. Stambler (2015) The Critical Need to Promote Research of Aging and Aging-related Diseases to Improve Health and Longevity of the Elderly Population. *Aging and Disease*, 6, 1-5.
- Kim, S. J., J. Y. Sung, J. W. Um, N. Hattori, Y. Mizuno, K. Tanaka, S. R. Paik, J. Kim & K. C. Chung (2003) Parkin cleaves intracellular alpha-synuclein inclusions via the activation of calpain. *J Biol Chem*, 278, 41890-9.
- Klamroth, S., S. Steib, S. Devan & K. Pfeifer (2016) Effects of exercise therapy on postural instability in Parkinson disease: a meta-analysis. *Journal of Neurologic Physical Therapy*, 40, 3-14.
- Knapp, M., M. Prince, E. Albanese, S. Banerjee, S. Dhanasiri, J. L. Fernandez, C. Ferri, T. Snell & R. Stewart (2007) Dementia UK: report to the Alzheimer's Society. *Kings College London and London School of Economics and Political Science*.
- Konrad, P. (2006) The ABC of EMG A Practical Introduction to Kinesiological Electromyography. Version 1.4. *Noraxon INC [Электронный ресурс].-URL: <http://www.noraxon.com/docs/education/abc-of-emg.pdf>*.
- Kontis, V., J. E. Bennett, C. D. Mathers, G. Li, K. Foreman & M. Ezzati (2017) Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *The Lancet*, 389, 1323-1335.
- Kotagal, V., R. L. Albin, M. L. T. M. Müller, R. A. Koeppe, K. A. Frey & N. I. Bohnen (2013) Diabetes is associated with postural instability and gait difficulty in Parkinson disease. *Parkinsonism & related disorders*, 19, 522-526.
- Koza, J. R. 1992. *Genetic programming: on the programming of computers by means of natural selection*. MIT press.
- Lacy, S. E., M. A. Lones, S. L. Smith, J. E. Alty, D. R. Jamieson, K. L. Possin & N. Schuff. 2013. Characterisation of movement disorder in parkinson's disease using evolutionary algorithms. 1479-1486. ACM.
- Lance, J. W., R. S. Schwab & E. A. Peterson (1963) Action tremor and the cogwheel phenomenon in Parkinson's disease. *Brain*, 86, 95-110.
- Langdon, W. B. & R. Poli. 2002. *Foundations of Genetic Programming*. Springer.
- Le Van Quyen, M. & A. Bragin (2007) Analysis of dynamic brain oscillations: methodological advances. *Trends in neurosciences*, 30, 365-373.
- LeCun, Y., Y. Bengio & G. Hinton (2015) Deep learning. *nature*, 521, 436.
- Lees, A. J. (2007) Unresolved issues relating to the shaking palsy on the celebration of James Parkinson's 250th birthday. *Mov Disord*, 22 Suppl 17, S327-34.
- Lilly, J. M. & S. C. Olhede (2009) Higher-Order Properties of Analytic Wavelets. *IEEE Transactions on Signal Processing*, 57, 146-160.

- (2010) On the analytic wavelet transform. *IEEE Transactions on Information Theory*, 56, 4135-4156.
- (2012) Generalized Morse wavelets as a superfamily of analytic wavelets. *IEEE Transactions on Signal Processing*, 60, 6036-6041.
- Litvan, I., J. G. Goldman, A. I. Tröster, B. A. Schmand, D. Weintraub, R. C. Petersen, B. Mollenhauer, C. H. Adler, K. Marder & C. H. Williams - Gray (2012) Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society Task Force guidelines. *Movement Disorders*, 27, 349-356.
- Liu, X., C. B. Carroll, S. Y. Wang, J. Zajicek & P. G. Bain (2005) Quantifying drug-induced dyskinesias in the arms using digitised spiral-drawing tasks. *J Neurosci Methods*, 144, 47-52.
- Lones, M. A., S. L. Smith, J. E. Alty, S. E. Lacy, K. L. Possin, D. R. S. Jamieson & A. M. Tyrrell (2014) Evolving Classifiers to Recognize the Movement Characteristics of Parkinson's Disease Patients. *Ieee Transactions on Evolutionary Computation*, 18, 559-576.
- Lopez, A. D. & C. C. J. L. Murray (1998) The global burden of disease, 1990–2020. *Nature medicine*, 4, 1241.
- Louis, S. J. & G. J. E. Rawlins. 1991a. Designer Genetic Algorithms: Genetic Algorithms in Structure Design. 53-60.
- . 1991b. *Using genetic algorithms to design structures*. Computer Science Department, Indiana University Bloomington, Indiana.
- Luo, J., K. Ying & J. Bai (2005) Savitzky–Golay smoothing and differentiation filter for even number data. *Signal Processing*, 85, 1429-1434.
- Macerollo, A., A. Batla, P. Kassavetis, I. Parees, K. P. Bhatia & M. J. Edwards (2015) Using reaction time and co-contraction to differentiate acquired (secondary) from functional 'fixed' dystonia. *J Neurol Neurosurg Psychiatry*, 86, 933-934.
- Maind, S. B. & P. Wankar (2014) Research paper on basic of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2, 96-100.
- Malfait, N. & T. D. Sanger (2007) Does dystonia always include co-contraction? A study of unconstrained reaching in children with primary and secondary dystonia. *Experimental brain research*, 176, 206-216.
- Mallat, S. 1999. *A wavelet tour of signal processing*. Academic press.
- Marteniuk, R. G., J. L. Leavitt, C. L. MacKenzie & S. Athenes (1990) Functional relationships between grasp and transport components in a prehension task. *Human Movement Science*, 9, 149-176.
- Martin, E. 2011. Novel method for stride length estimation with body area network accelerometers. 79-82. IEEE.
- Martin, W. E., R. B. Loewenson, J. A. Resch & A. B. Baker (1973) Parkinson's disease. Clinical analysis of 100 patients. *Neurology*, 23, 783-90.
- Mateus, C. & J. Coloma (2013) Health economics and cost of illness in Parkinson's disease. *European Neurological Review*, 8, 6-9.
- Matsumoto, S., M. Nishimura, H. Shibasaki & R. Kaji (2003) Epidemiology of primary dystonias in Japan: comparison with Western countries. *Movement disorders*, 18, 1196-1198.
- McCarthy, J. (1960) Recursive functions of symbolic expressions and their computation by machine, Part I. *Communications of the ACM*, 3, 184-195.
- Mehta, A. R., J. B. Rowe, M. R. Trimble, M. J. Edwards, K. P. Bhatia & A. E. Schrag (2013) Coactivation sign in fixed dystonia. *Parkinsonism & related disorders*, 19, 474.
- Merletti, R. & D. Farina (2009) Analysis of intramuscular electromyogram signals. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 367, 357-368.

- Miller, J. & P. Thomson. 2000. Cartesian Genetic Programming. In *Genetic Programming*, eds. R. Poli, W. Banzhaf, W. Langdon, J. Miller, P. Nordin & T. Fogarty, 121-132. Springer Berlin Heidelberg.
- Miller, J. F. (1999) An empirical study of the efficiency of learning boolean functions using a cartesian genetic programming approach. *Proceedings of the Genetic and Evolutionary Computation Conference*, 2, 1135-1142.
- (2011a) Cartesian genetic programming. *Cartesian Genetic Programming*, 4-14.
- (2011b) Cartesian genetic programming. *Cartesian Genetic Programming*, 17-34.
- . 2013. Gecco 2013 tutorial: Cartesian genetic programming. 715-740. ACM.
- Miller, J. F. & S. L. Smith (2006) Redundancy and computational efficiency in cartesian genetic programming. *IEEE Transactions on Evolutionary Computation*, 10, 167-174.
- Morgante, F., M. J. Edwards, A. J. Espay, A. Fasano, P. Mir & D. Martino (2012) Diagnostic agreement in patients with psychogenic movement disorders. *Movement Disorders*, 27, 548-552.
- Morris, J. C. (1997) Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International psychogeriatrics*, 9, 173-176.
- Mortazavi, D., A. Z. Kouzani & H. Soltanian-Zadeh (2012) Segmentation of multiple sclerosis lesions in MR images: a review. *Neuroradiology*, 54, 299-320.
- Murray, C. J., A. D. Lopez & D. T. Jamison (1994) The global burden of disease in 1990: summary results, sensitivity analysis and future directions. *Bulletin of the World Health Organization*, 72, 495.
- Murray, C. J. L., A. D. Lopez & O. World Health (1996) The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary.
- Nasreddine, Z. S., N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings & H. Chertkow (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53, 695-699.
- Newby, R. E., D. E. Thorpe, P. A. Kempster & J. E. Alty (2017) A History of Dystonia: Ancient to Modern. *Movement Disorders Clinical Practice*, 4, 478-485.
- Neychev, V. K., X. Fan, V. I. Mitev, E. J. Hess & H. A. Jinnah (2008) The basal ganglia and cerebellum interact in the expression of dystonic movement. *Brain*, 131, 2499-2509.
- Ngui, W. K., M. S. Leong, L. M. Hee & A. M. Abdelrhman. 2013. Wavelet analysis: Mother wavelet selection methods. 953-958. Trans Tech Publ.
- Nussbaum, R. L. & C. E. Ellis (2003) Alzheimer's Disease and Parkinson's Disease. *New England Journal of Medicine*, 348, 1356-1364.
- Nutt, J. G., M. D. Muentner, A. Aronson, L. T. Kurland & L. J. Melton, 3rd (1988) Epidemiology of focal and generalized dystonia in Rochester, Minnesota. *Mov Disord*, 3, 188-94.
- O'Neill, M. & C. Ryan (2001) Grammatical evolution. *IEEE Transactions on Evolutionary Computation*, 5, 349-358.
- Olhede, S. C. & A. T. Walden (2002) Generalized morse wavelets. *IEEE Transactions on Signal Processing*, 50, 2661-2670.
- Onwubolu, G. C. & B. V. Babu. 2013. *New optimization techniques in engineering*. Springer.
- Oppenheim, H. (1911) Über eine eigenartige Krampfkrankheit des kindlichen und jugendlichen Alters (Dysbasia lordotica lordotica progressive, Dystonia musculorum deformans). *Neurol Cbl*, 30, 1090-1107.
- Orru, G., W. Pettersson-Yeo, A. F. Marquand, G. Sartori & A. Mechelli (2012) Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, 36, 1140-1152.

- Owan, Y., H. Murakami, Y. Mori, K. Yamagishi, D. Watanabe, H. Kato, M. Kezuka & M. Kawamura (2015) Correlation between cognitive impairment and postural instability in patients with Parkinson's disease. *Brain and nerve= Shinkei kenkyu no shinpo*, 67, 99-104.
- Pal, P. K. (2011) Electrophysiologic Evaluation of Psychogenic Movement Disorders. *Journal of movement disorders*, 4, 21.
- Pandey, S., F. Nahab, J. Aldred, J. Nutt & M. Hallett (2014) Post - Traumatic Shoulder Movement Disorders: A Challenging Differential Diagnosis Between Organic and Functional. *Movement disorders clinical practice*, 1, 102-105.
- Parent, M. & A. Parent (2010) Substantia nigra and Parkinson's disease: a brief history of their long and intimate relationship. *Can J Neurol Sci*, 37, 313-9.
- Pareés, I., M. Kojovic, C. Pires, I. Rubio-Agusti, T. A. Saifee, A. Sadnicka, P. Kassavetis, A. Macerollo, K. P. Bhatia & A. Carson (2014) Physical precipitating factors in functional movement disorders. *Journal of the neurological sciences*, 338, 174-177.
- Parkinson, J. 2002. An essay on the shaking palsy. 1817. 223–36; discussion 222. *J. Neuropsychiatry Clin. Neurosci.* **14** (2).
- Parkinson's UK. <https://www.gov.im/media/1355777/freezing-in-parkinsons.pdf>.
- Paul, T. K. & H. Iba (2009) Prediction of cancer class with majority voting genetic programming classifier using gene expression data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 6, 353-367.
- Pearce, J. (2005) A note on scrivener's palsy. *Journal of Neurology, Neurosurgery, and Psychiatry*, 76, 513-513.
- Pedersen, K. F., J. P. Larsen, O.-B. Tysnes & G. Alves (2013) Prognosis of mild cognitive impairment in early Parkinson disease: the Norwegian ParkWest study. *JAMA neurology*, 70, 580-586.
- Pegasus 3D mesh. 2017. <http://web3d.jondgoodwin.com/recipes/mesh.html>.
- Perez, F., C. Helmer, A. Foubert-Samier, S. Auriacombe, J.-F. Dartigues & F. Tison (2012) Risk of dementia in an elderly population of Parkinson's disease patients: a 15-year population-based study. *Alzheimer's & Dementia*, 8, 463-469.
- Picardi, C., J. Cosgrove, S. L. Smith, S. Jamieson & J. E. Alty. 2017. Objective Assessment of Cognitive Impairment in Parkinson's Disease Using Evolutionary Algorithm. 109-124. Springer.
- Playfer, J. R. (1997) Parkinson's disease. *Postgraduate Medical Journal*, 73, 257-264.
- Poli, R. 1996. *Parallel distributed genetic programming*. University of Birmingham, Cognitive Science Research Centre.
- . 1997. Evolution of Graph-Like Programs with Parallel Distributed Genetic Programming. 346-353.
- Poli, R., W. B. Langdon & N. F. McPhee. 2008. A Field Guide to Genetic Programming. lulu. com. ISBN 978-1-4092-0073-4.
- Poularikas, A. D. 2010. *Transforms and applications handbook*. CRC press.
- Prince, M., M. Knapp, M. Guerchet, P. McCrone, M. Prina, A. Comas-Herrera, R. Wittenberg, B. Adelaja, B. Hu & D. King (2014) Dementia UK: update. *Alzheimer's Society*.
- Qiao, L., A. Esmaeily & H. G. Melhem (2012) Signal pattern recognition for damage diagnosis in structures. *Computer - Aided Civil and Infrastructure Engineering*, 27, 699-710.
- Rajput, A. H., B. Rozdilsky & A. Rajput (1991) Accuracy of clinical diagnosis in parkinsonism--a prospective study. *Can J Neurol Sci*, 18, 275-8.
- Ramazzini, B. (2001) De Morbis Artificum Diatriba [Diseases of Workers]. *American Journal of Public Health*, 91, 1380-1382.
- Reaz, M. B., M. S. Hussain & F. Mohd-Yasin (2006) Techniques of EMG signal analysis: detection, processing, classification and applications. *Biological procedures online*, 8, 11-35.

- Reeve, A., E. Simcox & D. Turnbull (2014) Ageing and Parkinson's disease: why is advancing age the biggest risk factor? *Ageing Res Rev*, 14, 19-30.
- Rioul, O. & M. Vetterli (1991) Wavelets and signal processing. *IEEE signal processing magazine*, 8, 14-38.
- Rushworth, G. (1960) Spasticity and rigidity: an experimental study and review. *Journal of neurology, neurosurgery, and psychiatry*, 23, 99.
- Russakovsky, O., J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla & M. Bernstein (2015) Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115, 211-252.
- Ryan, C., J. J. Collins & M. O. Neill. Grammatical evolution: Evolving programs for an arbitrary language. 83-96. Springer.
- Scarpa, M. & U. Castiello (1994) Perturbation of a prehension movement in Parkinson's disease. *Movement disorders*, 9, 415-425.
- Scheper, R. A. & A. Teolis (2003) Cramer-Rao bounds for wavelet transform-based instantaneous frequency estimates. *IEEE Transactions on Signal Processing*, 51, 1593-1603.
- Schettino, L. F., S. V. Adamovich & H. Poizner (2003) Effects of object shape and visual feedback on hand configuration during grasping. *Experimental Brain Research*, 151, 158-166.
- Schettino, L. F., V. Rajaraman, D. Jack, S. V. Adamovich, J. Sage & H. Poizner (2004) Deficits in the evolution of hand preshaping in Parkinson's disease. *Neuropsychologia*, 42, 82-94.
- Schneider, K. & R. F. Zernicke (1989) Jerk-cost modulations during the practice of rapid arm movements. *Biological Cybernetics*, 60, 221-230.
- Schrag, A. E., K. P. Mehta Ar Fau - Bhatia, R. J. Bhatia Kp Fau - Brown, R. S. J. Brown Rj Fau - Frackowiak, M. R. Frackowiak Rs Fau - Trimble, N. S. Trimble Mr Fau - Ward, J. B. Ward Ns Fau - Rowe & J. B. Rowe The functional neuroimaging correlates of psychogenic versus organic dystonia.
- Schwefel, H.-P. (1977) Numerische Optimierung von Computer-Modellen (PhD thesis). *Reprinted by Birkhuser*.
- Segawa, M., A. Hosaka, F. Miyagawa, Y. Nomura & H. Imai (1976) Hereditary progressive dystonia with marked diurnal fluctuation. *Adv Neurol*, 14, 215-33.
- Selesnick, I. W., R. G. Baraniuk & N. C. Kingsbury (2005) The dual-tree complex wavelet transform. *IEEE signal processing magazine*, 22, 123-151.
- Shimmer3 ECG Unit. 2018. <http://www.shimmersensing.com/products/shimmer3-ecg-sensor>.
- Shulman, L. M., C. Singer, J. A. Bean & W. J. Weiner (1996) Internal tremor in patients with Parkinson's disease. *Mov Disord*, 11, 3-7.
- Siciliano, B. S., L. Villani, L. Oriolo, G. 2009. *Robotics Modelling, pp.106-107. Planning and Control*. London: Springer.
- Silva, S. & E. Costa (2009) Dynamic limits for bloat control in genetic programming and a review of past and current bloat theories. *Genetic Programming and Evolvable Machines*, 10, 141-179.
- Slavič, J., I. Simonovski & M. Boltežar (2003) Damping identification using a continuous wavelet transform: application to real data. *Journal of Sound and Vibration*, 262, 291-307.
- Smith, S. F. (1980) A learning system based on genetic adaptive algorithms.
- Smith, S. L., P. Gaughan, D. M. Halliday, Q. Ju, N. M. Aly & J. R. Playfer. 2007. Diagnosis of Parkinson's disease using evolutionary algorithms. 8:433-447. *Genet Program Evolvable Mach* DOI 10.1007/s10710-007-9043-9
- Spector, L. & A. Robinson (2002) Genetic programming and autoconstructive evolution with the push programming language. *Genetic Programming and Evolvable Machines*, 3, 7-40.

- Spelt, L., B. Andersson, J. Nilsson & R. Andersson (2012) Prognostic models for outcome following liver resection for colorectal cancer metastases: a systematic review. *European Journal of Surgical Oncology (EJSO)*, 38, 16-24.
- Stamelou, M., M. J. Edwards, M. Hallett & K. P. Bhatia (2011) The non-motor syndrome of primary dystonia: clinical and pathophysiological implications. *Brain*, 135, 1668-1681.
- Steeves, T. D., L. Day, J. Dykeman, N. Jette & T. Pringsheim (2012) The prevalence of primary dystonia: a systematic review and meta-analysis. *Mov Disord*, 27, 1789-96.
- Stojanović, M., D. Cvetković & V. S. Kostic (1995) A genetic study of idiopathic focal dystonias. *Journal of Neurology*, 242, 508-511.
- Szu, H., B. Telfer & J. Garcia (1996) Wavelet transforms and neural networks for compression and recognition. *Neural networks*, 9, 695-708.
- Teulings, H.-L., J. L. Contreras-Vidal, G. E. Stelmach & C. H. Adler (1997) Parkinsonism reduces coordination of fingers, wrist, and arm in fine motor control. *Experimental neurology*, 146, 159-170.
- The Dystonia Society. 2018a. <https://www.dystonia.org.uk/functional-dystonia>.
- . 2018b. <https://www.dystonia.org.uk/Pages/Category/dystonia-symptoms>.
- The National Institute of Neurological Disorders and Stroke. 2018. <https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Dystonias-Fact-Sheet>.
- The Open group. 2017. [http://pubs.opengroup.org/onlinepubs/9699919799/basedefs/V1\\_chap04.html#tag\\_04\\_16](http://pubs.opengroup.org/onlinepubs/9699919799/basedefs/V1_chap04.html#tag_04_16).
- The World Bank. 2003. World development indicators. Washington, DC.
- Torres-Russotto, D. & J. S. Perlmutter (2008) Task-specific Dystonias. *Annals of the New York Academy of Sciences*, 1142, 179-199.
- Tu, C.-L., W.-L. Hwang & J. Ho (2005) Analysis of singularities from modulus maxima of complex wavelets. *IEEE Transactions on Information Theory*, 51, 1049-1062.
- Turing, A. (2004) Intelligent machinery (1948). *B. Jack Copeland*, 395.
- Turner, A. & J. Miller. 2014. Cartesian Genetic Programming: Why No Bloat? In *Genetic Programming*, eds. M. Nicolau, K. Krawiec, M. Heywood, M. Castelli, P. García-Sánchez, J. Merelo, V. Rivas Santos & K. Sim, 222-233. Springer Berlin Heidelberg.
- van den Berg, C., P. J. Beek, R. C. Wagenaar & P. C. van Wieringen (2000) Coordination disorders in patients with Parkinson's disease: a study of paced rhythmic forearm movements. *Exp Brain Res*, 134, 174-86.
- van Erkel, A. R. & P. M. T. Pattynama (1998) Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. *European Journal of Radiology*, 27, 88-94.
- Walker, J. A., J. F. Miller & R. Cavill. 2006. A Multi-chromosome Approach to Standard and Embedded Cartesian Genetic Programming. 903-910.
- Wang, H. & G. Huang (2011) Application of support vector machine in cancer diagnosis. *Medical oncology*, 28, 613-618.
- Website of the parkinsons's disease society. 2016. <http://www.parkinsons.org.uk>.
- Williams, D. R., H. C. Watt & A. J. Lees (2006) Predictors of falls and fractures in bradykinetic rigid syndromes: a retrospective study. *J Neurol Neurosurg Psychiatry*, 77, 468-73.
- Williams-Gray, C. H., T. Foltynie, C. E. G. Brayne, T. W. Robbins & R. A. Barker (2007) Evolution of cognitive dysfunction in an incident Parkinson's disease cohort. *Brain*, 130, 1787-1798.
- Williams-Gray, C. H., S. L. Mason, J. R. Evans, T. Foltynie, C. Brayne, T. W. Robbins & R. A. Barker (2013) The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. *J Neurol Neurosurg Psychiatry*, jnnp-2013.

- Wimo, A., L. Jönsson, J. Bond, M. Prince & B. Winblad (2013) The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 9, 1-11.e3.
- Winkler, S. M., M. Affenzeller & S. Wagner (2009) Using enhanced genetic programming techniques for evolving classifiers in the context of medical diagnosis. *Genetic Programming and Evolvable Machines*, 10, 111-140.
- workout trends. 2013. <http://workouttrends.com/arm-muscles-and-workout>.
- World Health, O. 2006. *Neurological disorders: public health challenges*. World Health Organization.
- Yarnall, A. J., D. P. Breen, G. W. Duncan, T. K. Khoo, S. Y. Coleman, M. J. Firbank, C. Nombela, S. Winder-Rhodes, J. R. Evans & J. B. Rowe (2014) Characterizing mild cognitive impairment in incident Parkinson disease The ICICLE-PD Study. *Neurology*, 82, 308-316.
- Zebra press. 2014. <https://thaw.org/tag/authentication/>.
- Zhang, J., J. Xu, X. Hu, Q. Chen, L. Tu, J. Huang & J. Cui (2017) Diagnostic method of diabetes based on support vector machine and tongue images. *BioMed research international*, 2017.
- Zhang, M. & P. Wong (2008) Genetic programming for medical classification: a program simplification approach. *Genetic Programming and Evolvable Machines*, 9, 229-255.
- Zhongshen, L. (2003) The design of butterworth lowpass filter based on MATLAB [J]. *Heilongjiang Electronic Technology*, 3, 017.