Imaging and Data Analysis as a Support for Diagnosis and Rehabilitation Practice

Candidate: Marco Trombini

Supervisor: Prof. Silvana Dellepiane

Coordinator of the PhD Course: Prof. Maurizio Valle

A thesis presented for the degree of
Doctor of Philosophy in Science and Technology for
Electronic and Telecommunications Engineering

Cycle XXXIV (2018-2021)

Department of Electric, Electronic, and Telecommunication
Engineering and Naval Architecture - DITEN
Università degli Studi di Genova

b

Preface

The present thesis is prepared in fulfillment of the requirements for acquiring the degree of Doctor of Philosophy in Science and Technology for Electronic and Telecommunications Engineering.

The PhD in Science and Technology for Electronic and Telecommunications Engineering (STIET), is promoted by the Department of Electrical, Electronic, and Telecommunication Engineering and Naval Architecture (DITEN) - Università degli Studi di Genova, in agreement with the Italian Institute of Technology (IIT).

The candidate is pursuing the aforementioned degree with a curriculum in Electromagnetism, Electronics, Telecommunications. The research theme is *Web-based platforms for medical imaging and surgical and rehabilitation therapy for musculoskeletal system pathologies.*

The PhD grant is funded by Programma Operativo Regione Liguria FSE 2014-2020 (Asse 3 “Istruzione e Formazione”).

The presented activities aim at promoting the research on innovation and on digital and enabling technologies, supporting the enhancement of human capital, as a determining factor for the development of research and innovation in Italy.

The research activity is compliant with the National Strategy of Intelligent Specialization (*Strategia Nazionale di Specializzazione Intelligente* - SNSI) and the National Research Program (*Programma Nazionale per la Ricerca* - PNR), and coherent with Law 240/2010, with the aim of favoring innovation and interchange between the world of research and the world of production and the qualification of the contribution of the activities research in the fields of innovation (Law 240/2010, art.24, co. 3 and subsequent amendments).

Abstract

In the last years, medical technologies have acquired a fundamental role in the clinical practice. Indeed, they are employed in all the stages of patients' treatment, ranging from diagnosis to rehabilitation.

Medical imaging is an umbrella term covering several medical technologies for diagnosis. Images and volumes are useful in the analysis of tissues and anatomical structures, enabling an accurate understanding of the phenomena and a quantitative dynamic analysis associated with longitudinal studies.

In addition, Information and Communication Technologies (ICT) are a concrete answer to many needs in the healthcare framework since they enable remote support for patients, thus fostering continuity of care and personalized rehabilitation plans.

The aim of the present work is the development and use of medical technologies that take into account the global patient's medical journey. In particular, the focus is on both diagnosis and rehabilitation, as indeed they are strictly connected procedures.

The two research themes are disclosed in parallel. After the technical disclosure, their deployment into real scenarios is described, by referring to particular diseases.

The integration of medical technologies for diagnosis and rehabilitation is straightforward, and thus a general pipeline combining them is proposed as future development of the present work.

Acknowledgements

Once at the end of a journey it is time to express gratitude to the people who shared parts of it with you.

My very first appreciation goes to Professor Silvana Dellepiane, for driving me throughout the PhD, for sharing her knowledge and teaching me, for showing enthusiasm and always having my back, thus acting as mentor, supervisor, but also as a mate.

I would like to sincerely acknowledge my colleagues from DITEN Department, Professors, Researchers and PhD Students, for their wonderful collaboration and fellowship. Also, I would like to thank all the people I collaborated with, for the fruitful discussions and competence exchange. In addition, I want to thank all the persons that helped me out though the years for reaching this goal and developing my professional background.

My deepest gratitude goes to my parents, my brother, and my girlfriend, for always supporting me in any choice of mine, for backing me up even in hard times, for their patience and perseverance in dealing with me, and for making me sure they can always be the rock I can safely rest on.

In the end, I want to thank all my friends and my family, for being a so important part of my life.

The end of a journey yields to the beginning of a new one, and the persons I met so far make me wish for the brightest future I could hope for, as long I am surrounded by people like them.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

In the biomedical context, diagnosis and rehabilitation are strictly connected procedures. Indeed, the purpose of diagnosis is the acquisition of sufficient information to understand the cause of the presenting problem and drive subsequent interventions [1], included the rehabilitation process. In addition, the earlier a disease is diagnosed, the more effective are cures and remedies for its treatment.

A widely used technique for diagnosis is medical imaging. Among the most important diagnostic imaging modalities one can mention radiography, X-ray Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Ultrasound Imaging (US) [2]. All of them have shown great potential for the early diagnosis of a wide spectrum of pathologies. At the same time, the recent development of novel data analysis and multidimensional visualization techniques improves the information available to the physicians during the diagnostic and follow-up process. Images and volumes are useful in the analysis of tissues and anatomical structures, enabling an accurate understanding of the phenomena and a quantitative dynamic analysis associated with longitudinal studies.

Once a disease is diagnosed, it is necessary to put into practice some medical therapy which, in several cases ranging from cognitive to motor diseases, includes rehabilitation. The clinical staff must consider that pathologies may rapidly evolve and that different patients may exhibit different manifestation of the disease. This means that the rehabilitation process must be personalized to each person, in order to cater their specific needs. In addition, the continuity of care is a crucial point for the treatment to be effective. In such a framework, the technological innovation makes it possible to meet the aforementioned requirements. Indeed, Information and Communication Technologies (ICT) and Internet of Things (IoT) solutions are becoming a concrete answer to many needs in the healthcare framework since they enable remote support for patients, thus fostering continuity of care and personalized rehabilitation plans.

The research project underlying the present thesis concerns a translational approach integrating studies and experiences of the Polytechnic and Medicine Schools. The goal is to improve patient management and support from the diagnosis to the rehabilitation phases.

Unfortunately, the spread of the Covid-19 pandemic, which impacted on about 2/3 of the whole PhD period, made difficult the access to medical facilities and data. As a results, it did not allow for focusing on a single disease from the diagnosis to the rehabilitation treatment. Conversely, several pathologies have been investigated, in order to exploit the two research themes in parallel and enabling the study of the full medical process as a future development.

The aims of the present work are:

- the development of novel techniques for signal and image processing, with particular focus on segmentation of Magnetic Resonance (MR) images and volumes;
- the application of such novel techniques to different pathologies, disclosing them in real-world scenarios;
- the employment of the telerehabilitation system ReMoVES [3] for treating a wide spectrum of diseases;
- the study of rehabilitation data via State of the Art (SoA) methods for data analysis;
- the definition of a general pipeline for the medical practice based on the proposed techniques.

In Chapter 2, the SoA revision is presented, with respect to both the two research themes. In particular, a brief introduction to diagnostic imaging, mainly focused on MRI, is disclosed, along with the image processing techniques that are currently used and the issues that need to be faced. Then, the context of ICT solutions in healthcare is introduced, focusing on rehabilitation and assistive technologies. To conclude, the diseases matter of application of the thesis are presented, also from a clinical point of view.

Chapters 3 and 4 describe in depth the methods and materials developed and involved in the PhD course, i.e., the main contributions of the present work, in diagnostic imaging and telerehabilitation respectively. Their application to real-world scenarios is disclosed in Chapters 5 and 6 respectively, focusing on the diseases introduced in Chapter 2.

To conclude, Chapter 7 provides a glimpse into future developments of the present study.

CHAPTER 2

State of the art

2.1 An introduction to diagnostic imaging

As the Norwegian playwright and theater director Henrik Ibsen first said, *"A thousand words leave not the same deep impression as does a single deed"*. Such a quote was then paraphrased into the famous saying *"A picture is worth a thousand words"*, which is valid in several fields, including the context of medicine.

For centuries, clinicians have been hindered the application of the most efficient human sense, i.e., vision, for diagnostic purposes on internal structure of the human body. However, the importance of images in the biomedical context is unquestionable. Images convey visual information about some property of the human body and are produced by different types of a scanner utilizing different physical phenomenon. For instance, CT utilizes the absorption abilities of x-rays of tissues, US utilizes the reflections of sound waves in the anatomical structures, MRI utilizes magnetic relaxation properties of hydrogen particles.

In general, a common scheme for producing medical images can be defined, consisting of six steps:

1. define the property to be imaged, based on the medical need.
2. Select a physical phenomenon to track and measure the aforementioned property.
3. Develop a mathematical model for quantifying the mentioned property.
4. Develop an algorithm for reconstructing the image based on the mapped property.
5. Build the imaging device.
6. Apply some signal and image processing for extracting information from the obtained data.

As for the present thesis, the focus is set on the last operational step. The type of image under analysis consistently affect the processing to be performed, hence, the fundamentals of the imaging techniques need to be disclosed. The several imaging modalities existing present different peculiarities and applications, thus providing clinicians with a variety of options to perform efficient and reliable examinations. In the next Sections, the main medical imaging modalities are introduced, with particular emphasis on MR.

Medical imaging techniques

When speaking of diagnostic imaging, some "major" modalities can be identified, i.e., the most used ones. X-rays and CT are classical examples. Here, the underlining procedure consists of irradiating the body with an x-ray beam. Different tissues interact differently with the x-ray photons. This phenomenon is due to differences in the electron density of each tissue type. Therefore, some radiation is absorbed or attenuated by the body itself, and some of the radiation passes through. A map of the attenuation can be obtained by placing a receiver behind the body, resulting in a two-dimensional image of the inner structure of the scanned body.

Another very important imaging modality is US, which takes its origin from the military arena. US is the most cost-effective medical imaging modality available today. The information content which is delivered by US is versatile as it enables measurements in 1D, 2D, 3D, and also 4D (space and time). The physical phenomena which US is based on is mechanical waves.

Among the other imaging techniques, it is worth mentioning Positron Emission Tomography (PET), Electrical Impedance Tomography (EIT), and Optical Coherence Tomography (OCT).

In the present discussion, MRI was left apart. Of course this is not a lapse of memory, but rather, such an imaging modality deserves a whole Section due to its central role in the underlining research project.

Magnetic Resonance Imagery

MRI is probably the most interdisciplinary imaging modality as it combines physics, mathematics, engineering, computer science, and last but not least, medicine. It is non-invasive and extremely versatile with numerous methods of distinguishing tissues with deep details. However, it is still rather expensive, the examination are rather time-demanding, and some issue in terms of of patient handling and accessibility may arise. For such reasons, the field of MRI is continuously growing and improving. A brief description of the fundamentals on MR follows.

The basic principle for MRI is the Nuclear Magnetic Resonance (NMR) phenomenon [4], wherein the nuclei of some atoms exhibit the capacity to absorb and re-emit radiofrequency energy when put in a magnetic field. Usually, the nuclei of hydrogen are considered, as they are widely spread in biological material and present good MR sensitivity. Such nuclei present a fundamental property which is called *spin*. Spin is an angular momentum, which, in standard conditions, is randomly directed. In MR, an external magnetic field, usually referred to as B_0, is applied, so that the direction of nuclei involved is aligned either to B_0 or $-B_0$. It hence possible to consider a *net magnetization vector* M, resulting from the difference between the parallel and anti-parallel aligned nuclei. The aim of MRI is to measure the modulus of M, in order to measure the concentration of nuclei in a region.

Spins deliver informative signals if the direction of the magnetization vector is altered. To this aim, a second magnetic field, usually called B_1 and perpendicular to B_0, is applied to excite the nuclei possessing spin. As a result, the net magnetization rotates of 90 degrees and spins around B_0. This can be detected as it induces an Alternating Current (AC) in the receiver coil specifically placed around the patient.

The excitation phase is then concluded and spins return to the equilibrium state. This causes the magnetization to decay over time, i.e., a decrease of the magnitude of M, thus leading to a decrease of the detected signal. The time required for the signal to return to equilibrium state is called *relaxation time*. Two particular relaxation types must be noted. The former, referred to as *longitudinal* *relaxation*, regards the process of realignment to B_0. Such a process is characterized by the so-called $T1$ *relaxation time*, i.e., the time needed by the system to recover 63% of its equilibrium value after the excitation, and has specific values for different types of anatomical structures. The latter relaxation type is the *transverse relaxation* and regards the spins precessing around the magnetization vector. Spin-spin interactions yield small random local magnetic inhomogeneities, resulting in different ways of dephasing. The time for this process is called $T2$ *relaxation time* and corresponds to the time needed by the dephasing to lead the signal to 37% of its original value. As T_1, T_2 depends on tissues, but is shorter in general.

Aimed at generating an image, spatial location of signal must be taken into account. For such purpose, a magnetic gradient is added along the main magnetic field, in the cranial-caudal direction. In this fashion, it is possible to select thin slices of spins through the body, and change slice by operating on the excitation phase. Then, *frequency* and *phase encoding* enable the identification of specific points in slices, i.e., pixels. Phase encoding consists of a small change applied in the field, influencing the frequency of precessing and resulting in a shift in the precessing phase. Then, frequency encoding is performed by applying a gradient during the readout phase of the signal, yielding to a specific shift of the resonance frequency and allowing the identification of pixels inside a specific phase encoding. Phase and frequency combinations identify specific pixels, which are placed on a grid, called *K-space*. To conclude, data are mapped into the intensity values of pixels via a Fast Fourier Transform (FFT).

MRI sequences: T1-weighted and STIR

The final appearance of MR images can be adjusted to specific clinical needs by applying gradients with different criteria. In particular, one can speak of MRI *sequences* as sets of gradients resulting in a set of images with a particular appearance. As the multitude of sequences available on modern scanners is large, in the research activity described in the present thesis only two have been considered, i.e., *T1-weighted* and Short Tau Inversion Recovery (STIR) [5]. The former produces images that are the closest approximation of the macroscopic appearance of anatomical structures, and are included in almost all MRI protocols. They are characterized by low signal intensity for fluid, mild signal intensity for muscle, an high signal intensity for fat. In the present thesis, T1-weighted MR images are used for segmentation purposes and refer to the anatomical district of the wrist. The latter sequence is aimed at suppressing the bright signal from fat. This is performed for instance after the administration of a contrast medium, to enhance the visibility of details otherwise not appreciable. In the present thesis, STIR images are considered for detection purposes related to an inflammatory rheumatic disease.

Image segmentation and ROIs detection

The analysis and interpretation of digital images is consistently aided by their partition into homogeneous areas, possibly corresponding to meaningful Regions of Interest (ROIs) in the scene. Such a partition process is called image segmentation. In particular, according to its classical definition [6], the segmentation result is the partition of an image into disjoint, non-empty, and connected subregions, for which some predicate of homogeneity is satisfied. In addition, the same predicate of homogeneity must not be valid for the union of any such subregions. As a natural development of such a definition, the purpose of segmentation can often be oriented to the detection of image regions that are meaningful within a particular application domain.

In the medical field, this fundamental activity is often performed manually [7], resulting in time consuming endeavors. Indeed, also enforced by the recent growth in data availability, automatic image segmentation techniques have gained a fundamental role in the image processing domain [8].

SoA segmentation techniques can be divided into deformable and parametric models. The methods in both categories have many advantages in terms of usability, as they can be completely unsupervised or they may not require huge amount of data.

Deformable models take origin from [9] and [10], and are defined by curves or surfaces that can move under the influence of internal forces coming within the model itself and external forces computed from the image data. Later, an ensemble of their geometrical features gave birth to models such as [11], [12], and [13]. They are still commonly used in multiple domains, as for example in the biomedical field [14], [15], and in remote sensing [16], [17].

Conversely, when dealing with parametric models, the segmentation problem is often formulated in terms of a Bayesian probabilistic graphical model [18]. The Bayesian statistical framework allows exploiting data and prior distributions [19], aimed at linking image segmentation with ROI detection [20]. In addition, it also fosters weakly supervised approaches [21], which take origin from saliency detection models [22].

In the present thesis, a method combining such a parametric framework with a graph-based formulation is presented. Indeed, graph-based approaches have been widely used in the signal processing context [23]. For instance, images (and volumes) can be associated with weighted graphs via the one-to-one correspondence between the pixels (or voxels) and the nodes of the graphs. Relations and similarity distances between pixels are generally expressed in terms of minimum-length path between nodes. As an example, the method described in [24], starts from a set of points, called seeds, and it looks for maximally connected paths and assigns a cost to each pixel, related to each seed. Such a procedure is equivalent to the fuzzy connectivity computation proposed in [25], which integrates intensity similarity with topological connectedness. In [24], a graph cut is performed after the cost-computation phase, in order to obtain the segmentation. One might prove that such a method can be formally described in terms of the Image-Foresting Transform (IFT) approach [26]. In [27], the IFT has been shown to be equivalent to the Minimum-Spanning Forest (MSF) cuts constrained by seeds, hence, some analogies between the graph-cut in [24] and other methods e.g., the threshold in the watershed hierarchy in [28], might be investigated.

A graph-based approach was also proposed in [29], where the IFT transform was extended to support superpixel [30] computation. However, according to the work in [31], the generation of superpixels is the result of an over-segmentation of the input image. As a consequence, the predicate of homogeneity is satisfied also by the union of subregions, thus being in contrast with the image segmentation definition. Indeed, the detection of meaningful image parts requires a further merging step.

Besides the aforementioned techniques, the availability of wide image datasets and the recent development in the deep learning (DL) literature [32] have yielded effective results for semantic segmentation, which is actually different with respect to classic image segmentation, because it aims at partitioning the image to obtain a class labeling and with the use of a training set. The main drawback of deep learning solutions is their need for very large labeled datasets. As a consequence, the concept of transfer learning (TL) has been developed [33]. In case smaller datasets are available it is possible to reuse and refine popular pre-trained networks, such as the one reported in [34], [35], [36], and [37].

Tackling human bias in automatic medical image analysis

Medical image analysis is a major interest in biomedical research [38]. Among the several topics tackled, one can mention segmentation, visualization, quantification, object tracking, and detection [39], which are strictly linked with the application for medical image diagnosis, therapy planning, and other quantitative studies.

Evaluation and validation of new methods for medical image analysis have been based on the authors' personal data sets for a long time, so that unbiased and reliable comparisons of the solution were hard to be performed [40].

Nowadays, international challenges have become the standard for validation of medical image analysis methods. However, the lack of quality control on such challenges may have substantial consequences in terms of reproducibility and interpretation of the results [41]. Also, the large variability in challenge design may affect the reliability of the validation. Indeed, several issues have been proved [41], such as:

- minor changes in metrics may make the last the first.
- Different data aggregation methods produce different winners.
- Different annotators produce different winners.
- Removing one test case can change the winner.

In such a context, the need for finding better ways to collect, annotate, and reuse medical imaging data is clear. The first key point for such a problem is that the classification label of each image data must be as accurate and reproducible as possible [42].

Building a robust dataset is much more difficult in the medical domain than in other more generic applications. Indeed, when referring to tasks such as detecting transports or traffic lights in photographic images, a large group of people can provide reliable annotations, as neither specific education nor particular experience is required. Therefore, also using some web-based tool, the pool of annotators is much larger and thus widespread annotations can be done.

Conversely, in the medical domain, supervised ground-truth images are provided by medical experts and present several issues, in terms of time for the procedure, raters reliability, and bias due to experts' knowledge.

Aimed at overcoming these problems and at fostering reliable medical image analysis, some operations are suggested, which may be useful either for standardizing the reference data or for making the evaluation process objective and independent on the raters. They are introduced with regards to a specific application, which will be disclosed in the corresponding section.

2.2 Technologies in rehabilitation

The ongoing digital transformation in our society has significant impact on several technological aspects, so that the term Fourth Industrial Revolution has been used for a few years. As another revolution [43], IoT solutions are becoming increasingly relevant, and their use is consistently growing in several application domains. Regarding healthcare, the IoT market size was valued at USD 147.1 billion in 2018 [44], and is expected to reach USD 534.3 billion by 2025, expanding at a compound annual growth rate (CAGR) of 19.9% over the forecast period [45], and resulting in an estimated USD 63 billion of savings due to the deployment of medical IoT by 2022 [46]. All this is due to growing investments in digital technology implementation at healthcare institutions that address the need for the care of a growing geriatric population [47] coupled with the rising prevalence of chronic conditions [48]. In addition, the recent outbreak of Covid-19 has had strong impact on the health system, which had to adapt itself to various needs such as guaranteeing access to care for patients in forced quarantine or in solitary confinement, and meeting the needs for social distancing and reduction in access to healthcare facilities. Medical IoT solutions are an essential tool for responding to patient care needs under safe conditions. Hence, applications such as telemedicine, remote patient monitoring, and interactive medicine have a precise and crucial position in the fight against the coronavirus, such that several nations officially recommended their use [49].

The key benefit of the IoT in the medical domain is connected technology. Devices are used for assessing patients' conditions, and monitoring and supporting rehabilitation, so that a personalized plan of care can be defined and kept updated. This also fosters continuity of care, enabling a patient to be supervised by a multidisciplinary team even after dehospitalization. The most ubiquitous of such devices are wearable or robotic devices, for instance, smart bands for data collection related to some physical activity [50] or other wearables for motion analysis, which can be devoted to specific body-part rehabilitation (e.g., shoulders [51] and knees [52]). Even though a deep interest in such devices is manifested in the healthcare context, wearables, robotic devices, and devices based on smartphone interaction are not very suitable for the elderly population or for dehospitalized and disabled patients. Indeed, to fully exploit the potential of an IoT solution, patients should be able to deal with it autonomously; however, the presence of wearable devices or controllers means that some external support may be needed for such activities.

Among the several implications of using IoT/ICT solutions in the healthcare domain, one can mention clinical and practical consequences. Indeed, traditionally, a great portion of physical therapy, rehabilitation, and assessment was based on a clinician's observations and judgment. Nowadays, sensors and computing technologies can be used for motion capture, performance assessment, and range-of-movement (ROM) measurements. From a clinical point of view, patients benefit from the use of such technologies in terms of help for dehospitalization, continuity of care, the personalization of plans of care, and engagement in activities. In addition, from an operative point of view, telerehabilitation helps clinical staff to also follow several patients when they cannot physically attend to them. This is very important in the time of the pandemic emergency, as it, for instance, allows for reducing time for moving from one patient's house to the next. As a practical example, the Liguria region has relevant problems in terms of urban traffic; hence, home-based rehabilitation often causes therapists to lose time in traffic, augmenting work stress and eventually affecting the quality of the imparted treatment.

To conclude, healthcare technology is also contextualized in the framework of user-generated content (UGC) analysis. Therefore, it finds applications in research on the use of data science (DS) in digital marketing (DM). In 2020, nine topics for future research on DS in the DM ecosystem were defined [53]. Among them, four will be addressed in the present work, i.e., medical-data eHealth; people: movement, organization, and personalization; the IoT; and new machine-learning model development.

Exergame in rehabilitation

In the field of assistive technologies, exergames have recently gained great popularity and demonstrated scientific reliability, thus surpassing their original goal of mere entertainment. The term exergame refers to video games that impart physical exercise / support rehabilitation practice (in the context of their clinical application) in which the repetitive and task-oriented components of rehabilitation activities are reformulated in video-playful terms. There are various examples in the literature, e.g., of efficacy in maintenance and in improvement of cognitive functions in the elderly population [54] and in post-stroke rehabilitation [55]. The exergames can also be considered as a virtual reality (VR) tool, which can be a safe tool to access activities otherwise not accessible to the person with cognitive and motor disabilities in everyday life contexts. Furthermore, the *gamification* [56] determines a motivating and engaging environment in order to keep the patient busy without inducing boredom or fatigue, with consequent frustration and abandonment of therapy.

By combining this type of activity with a system capable of acquiring data and information during the exergame, it is possible to provide the therapist with a tool for an objective evaluation of performance, as well as to monitor and evaluate specific motor / cognitive tasks even in the absence of staff. clinical. The collection of data during the exercises is the necessary feature to give the therapist the opportunity to understand the patient's behavior, assess fatigue and possibly correct erroneous attitudes. In this way, the clinical staff maintain control over the patient's activity, follow progress and improvements and obtain the necessary foundation to prescribe a personalized care plan.

Data analytics and exergame

The spread of connected technologies delivering exergames fosters the access and collection of data. As a result, it enables the development of data analysis and artificial intelligence (AI) techniques, also for supporting clinical practice.

In such a context, a first application of DS regards gesture recognition. In [57] a yoga teacher performed a set of exercises which are recorded to construct a gestures training set. Based on that, students can practice such exercises in front of a Microsoft Kinect, and a trained algorithm automatically discerns correct and incorrect movements. In addition, a clinical application is reported in [58].

Among the machine learning (ML) techniques applied to rehabilitation evaluation, one can mention Dynamic Time Warping (DTW) [59]. It is a technique for comparing two-time series that can vary in speed and measure the similarity between them, in terms of minimum Euclidean distance. Its main application is for speech recognition but it can be employed also in gesture analysis, by expanding it from temporal to spatial domain. Comparisons between exergame data can be obtained also by using Hidden Markov Model (HMM) [60]. This is a stochastic approach that compute the probability that a sequence of observations (i.e., gestures and movements) are generated by the same process. Both these techniques are employed in [61].

Another approach for classifying gestures is support vector machine (SVM) [62]. Such a model can be used either for binary and multi-class classification, by switching to the *one-to-one* or *one-to-rest* formulation [63]. Some examples of work proposing such an approach are [64, 65, 66].

To conclude, it is worth noting that rehabilitation data are characterized by a repetitive and sequential nature. Hence, approaches such as long-short term memory (LSTM) [67] recurrent neural network (RNN) are straightforward to be used [68].

2.3 Application domains

In this section, a brief description of the diseases taken into account for the experimental phase is provided. They refer to different areas, including musculoskeletal and rheumatic diseases, cognitive diseases, and degenerative diseases.

Rheumatoid arthritis

Rheumatoid arthritis (RA) is a long-term autoimmune disorder mainly affecting joints and bones [7]. Bone erosions are referred to as the central feature of RA [69]. Erosions are marginal and localized at the “bare area”, i.e., the bone surface within the synovial space, which is not protected by cartilage [70], and are typically cause by the chronic inflammation. RA is consistently spread in the world, and about $0.5 - 1\%$ of a population suffers from such a disease [71]. Of course, it has also several economic implications, as it is estimated that over 1.5 million people in the USA suffer from RA, resulting in a cost of about 80 billion dollars for treatments per year [71]. RA manifests itself between ages 30 and 60 for women, later for men. Wrist joints are location mostly affected by RA [72].

According to the most recent recommendations, MRI is used to asses RA in terms of synovitis, bone edema, and bone erosions [71]. In particular, an MRI-based scoring system - Rheumatoid Arthritis Magnetic Resonance Imaging Studies (RAMRIS) - was developed, which is widely used by rheumatologists [73].

Of course, an early diagnosis of RA is essential for effective treatment of the disease. To this purpose, image processing techniques, such as segmentation, applied to MR images result very useful for diagnostic purposes in such a context, as they can make easier doctor's visualizations and enable details enhancement.

Spondyloarthitis

Axial spondyloarthritis (axSpA) is an inflammatory rheumatic disease, involving primarily an axial skeleton and progressively leading to the sacroiliac, intervertebral and facet joint immobilization [74]. Active sacroiliitis are identified on the basis of the Assessment in SpondyloArthritis international Society (ASAS) criteria [75], which aims at searching for bone marrow edema, that is visible for instance in STIR images [76]. Lesions are located periarticularly in a subchondral bone, and must be visible either on two consecutive slices of MRI examination, or on only one slice if at least two lesions are noticeable [77]. It is hence clear that the assessment of active sacroiliitis in MRI is not simple, and becomes even harder as lesions are small. As a result, inter-rater agreement based on ASAS criteria is still unsatisfactory [78].

Systemic sclerosis

Systemic Sclerosis (SSc) is a rare autoimmune rheumatic disease characterized by vascular injury, immune dysfunction, and an excessive production and accumulation of collagen, called fibrosis, that can affect the skin and internal organs including lungs, gastrointestinal tract and cardiovascular system [79] [80]. One of the major impairments, caused by skin induration and joint and muscle involvement, is the gradual loss of mobility which substantially affects the quality of life [81]. In particular, hand disabilities in SSc are frequent and contribute to the manifestation of diseases such as inflammatory arthritis, tendon friction rubs, tendonitis/tendinosis, puffy hands, skin sclerosis, calcinosis, acro-osteolysis, Raynaud's phenomenon and digital ulcers [82]. Furthermore, finger flexion and extension are the most impaired aspects of hand mobility in SSc patients [83]. Indeed, severe skin thickness in the hands can cause deformity in the flexion of the fingers, leading to the loss of flexion at the metacarpophalangeal (MCP) joints, the loss of extension of the proximal interphalangeal (PIP) joints, and the loss of thumb abduction. Moreover, the distal interphalangeal (DIP) joint may also become fixed in mid-range flexion. These impairments cause a claw-type deformity of MCP extension, PIP flexion, and thumb adduction.

As for skin ulcerations, they generally occur over joint contractures due to increased skin pressure in areas of bony prominences and reduced blood flow to the skin from scleroderma vasculopathy [82]. Focusing on the hands, small joint contractures yield dissatisfaction with appearance, social embarrassment and difficulties in carrying out work activities, thus resulting in a significant drawback for SSc patients [82] [83].

These various manifestations of hand impairment can result in reduced mobility, dexterity, and grip strength, recognize them is essential, although there is no definitive medical treatment options yet [82]. In patients with SSc, hand rehabilitation aims at improving hand mobility, functionality and strength as well as increasing involvement in daily living activities [81]. The role of rehabilitation treatment is crucial and involves a multidisciplinary team consisting of physicians, physiotherapists, and occupational therapists. Even though most therapists recognize the importance of rehabilitation for SSc, there is currently minimal awareness and SSc rehabilitation therapy is not widespread [83].

Unilateral spatial neglect

Unilateral spatial neglect (USN) is a term which denotes the failure to detect, respond or orient toward stimuli located in the hemibody and or hemispace contralateral to the lesioned hemisphere [84]. It is a common and severely disabling neurobehavioral disorder induced by discrete cerebral lesions such as stroke, tumors, trauma, and degenerative diseases. It is most frequently associated with damage involving the right parietal and occipital lobes, the basal ganglia and the thalamus [85]. Patients suffering from USN show slow functional progress during rehabilitation and need long hospitalization times, and are less likely to be able to live independently with deterioration of the quality of life. Three bedside tests can be used to detect the presence of USN: *line bisection*, *target cancellation*, and *drawing*. In line bisection, the patient is asked to mark the middle of twelve lines presented on a paper. The longer are the lines, the easier is to reveal neglect. With target cancellation, the patient is asked to mark out or cancel the targets (i.e., bells, stars or apples) among a variety of distractors. In particular, Apples cancellation test is aimed at differentiating between allocentric (object-centered) and egocentric (stimuli-centered) form of neglect [86]. Some other cancellation tasks, instead, have only target items (Albert's task) [87]. In drawing, the patient is required to copy a figure, or to draw it from memory.

Several rehabilitation strategies for USN (e.g., optokinetic stimulation, somatosensory electrostimulation, mirror therapy) have been reported [88]. Their efficacy, however, is still a matter of debate. A meta-analysis by Pollock et al. [89] states that there is only a limited number of high-quality studies suggesting the efficacy of USN interventions in improving functional outcomes and reducing disability. Azouvi et al. [90], in a recent review, concluded that there still is only a low level of evidence for the different rehabilitation methods and emphasized the need for longer validation trials using innovative techniques such as non-invasive brain stimulation (NIBS). NIBS techniques have recently emerged in restorative neurology due to their hypothetical advantage in enhancing the efficacy of traditional therapeutic intervention. In this view, the re-discovery of the application of a direct-current flow of low intensity (1 - $2 \mu A$) has raised much interest. This technique is known as transcranial direct-current stimulation (tDCS). It acts by a tonic modulation of the resting membrane potential of the cortical neurons, which occurs in an opposite direction, depending on the polarity (anodal vs cathodal) of the electrodes, placed on the chosen areas. It is commonly stated that cathodal stimulation (C-tDCS) decreases cortical excitability due to neural hyperpolarization, while anodal stimulation (A-tDCS) reaches the opposite effect by a subthreshold depolarization [91]. The use of tDCS has been shown to be a promising approach in order to improve post-stroke neglect. A-tDCS and bilateral tDCS appears to be more effective than C-tDCS [92]. A cognitive therapy is usually associated to the tDCS approach, which consists of performing tasks with the aim of improving the patient's capability of investigating the ignored hemispace. This complementary activity can be provided by digital solutions as in [93, 94, 95, 96, 97].

Post stroke

Prior the spread of the Covid-19 pandemic, stroke was usually referred to as the 21^{st} *Century epidemic* by the medical community. Stroke is indeed the second leading cause of death worldwide, and also the second most common cause of disability-adjusted life years (DALYs) [98]. Many people experiencing a stroke are left permanently disabled, placing a burden on family and society. It is then clear that stroke is an important sanitary emergency and, therefore, also the need for rehabilitation has become a very crucial issue.

In such a context, cerebellar stroke accounts for approximately 2% to 3% of all strokes. Acute cerebellar manifests itself with axial and or limb ataxia, nystagmus, vertigo, action tremor and dysarthria. The cerebellum works as a motor feedback control system: it compares the motor command elaborated in premotor areas with sensory-motor inputs, and then produces an error signal. A cerebellar damage can impair its ability to sufficiently integrate sensory input in order to monitor and correct movements. It is therefore possible to observe abnormalities in movement characterized by increased variability and poor accuracy. These disorders result in unsteady gait, increased postural sway, abnormal eye movements, uncoordinated limb movements, difficulties in speech.

Ataxia causes a general in-coordination of movement, in particular in dynamic ataxia there is a lack of control of distal segments of the body. For instance, equilibrium and walk are largely involved in studies as [99], [100], [101]. The study in [102] reported that cerebellar damage might cause an inertial mismatch between an internal representation of body dynamics and the actual body dynamics. According to these results, there are hypometric and hypermetric patients which respectively underestimate or overestimate their limb's inertia. This work also showed that altering the apparent inertia of the limb to correct the mismatch via robot could improve simple single-joint elbow movements for both types of patients. It follows that for single-joint movements, hypometric patients should theoretically improve with the addition of mass to the limb because it would reduce the discrepancy between the internal model and the actual limb dynamics. The research in [103] has already shown interest towards this peculiar strategy. Some immediate benefits of weighting on single-jointed elbow movements were highlighted, but no benefits on multi-jointed reaching movements were evident.

Frailty

Frailty in the elderly corresponds to a broad clinical issue that concerns the physical, cognitive, and social aspects of the patient, particularly for people over the age of 75 [104]. The study by Fried et al. [105] defined a phenotype and thereby some characteristic traits of frailty in the elderly. Specifically, frailty is considered if at least three of the following symptoms are present: unintentional weight loss, fatigue, reduction in muscle strength, slower walking speed, and decreased physical activity. In cases where fewer than 3 of the symptoms are detected, one can speak of prefrailty. Frailty, therefore, differs from disability because it is characterized by a decline in several physiological aspects. Thus, in this sense, disability manifests itself more as a consequence of frailty itself.

Due to age and related cognitive impairments, weakness is a major limiting factor related to daily life activities. For instance, the reduction in torque generation is reported at the level of the elbow, shoulder, fingers, and thumb, which worsens due to prolonged physical inactivity. Furthermore, simple activities, such as standing up, may be affected, causing falling risk and insecure gait. In addition to cases of psychiatric and neurological diseases, cognitive abilities inevitably decline in a healthy elderly population, thus leading to severe social and economic impact.

In this context, strength training associated with task-oriented training can intensify rehabilitation and reinforcement [106]. Th study of Erickson et al. [107] suggested that physical exercise can produce cognitive improvements (associated with an increase in hippocampal volume) in accordance to [108] about increased levels of brain-derived neurotrophic factor (BDNF) in response to exercise. By design, exergames are appropriate for this aim as they require the patient to produce physical movements in order to complete a task-oriented exercise in response to visual cues [109]. They are simultaneously able to improve patient engagement and train multiple cognitive processes [110].

Multiple sclerosis

Multiple Sclerosis (MS) is a chronic, inflammatory and demyelinating disorder of the central nervous system (CNS). It is characterized by a deficit of neurological functions, including motor, sensory, and cognition, which can be relapsing and/or progressive in nature [111]. Cognitive impairment (CI) can affect up to 70% of the MS population. Persons with MS can experience difficulties in several cognitive domains, including processing speed, sustained and selective attention, learning and episodic memory, with executive functions compromised in more advanced, progressive stages [112]. Other cognitive deficits such as visuospatial problems and difficulties with social functioning can also be present [113]. While more severe cognitive impairment is more likely in persons with secondary progressive MS, signs of cognitive involvement can be present early in the disease process. Within the first year of diagnosis, about half of persons with MS report having either minimal or mild cognitive difficulties, with greater complaints over the first decade. Although uncommon, some persons with MS present with cognitive impairment as their primary symptom. In addition, cognitive issues may be present preclinically [114]. These impairments correlate closely with brain pathology and can have major consequences for everyday life. Further, CI is the leading predictor of occupational disability in these patients. Treatment options remain extremely limited, however, despite an increase in available interventions. In principle, disease-modifying therapies might improve cognition in people with MS as these agents are primarily designed to arrest the disease and prevent relapses, but whether they directly improve cognition remains speculative [115].

Cognitive and behavioral rehabilitation strategies are designed to enhance an individual's capacity to process and interpret information and to function in all aspects of family and community life. Although the focus on designing and testing effective cognitive rehabilitation programs for people with MS is a relatively recent phenomenon, the growth in research studies addressing this need has been substantial over the past decade. As a result, the Italian National MS Society has recommended remedial interventions and accommodations that can be made to manage cognitive impairment and improve everyday functioning in both adult and pediatric MS populations. Such recommendations include more comprehensive assessment for anyone who tests positive for cognitive impairment on cognitive screening or demonstrates substantial cognitive decline, as well as neuropsychological evaluation for any unexplained change in academic performance or behavioral functioning in school-aged children with MS. Evidence suggests that cognitive rehabilitation has a long-term impact well beyond the treatment period and might enhance cognition in the face of future brain changes [116]. Such sustained effects have been documented in the literature on aging, in which cognitive rehabilitation not only improved everyday life activities, but also resulted in a 29% reduction in dementia risk ten years after treatment [116].

Balance

Balance is one of the most important ability for humans, as it is the principal responsible for postural control, and thus it affects the possibility of fulfilling or not countless activities, ranging from the most basic to harder ones [117]. The equation is simple, the more balance, the more postural control and thus less falls, risks of injury and more independence on activities of daily life. Despite this naive motivation on why balance is crucial, a myriad of factors contribute on the preservation of a sufficient balance in persons. Even though some of them may depend on external causes, several ones regard the health conditions of people. Indeed, one can find plenty of works in the recent literature studying balance and its relation with other diseases, with particular emphasis on the role of technology for favoring rehabilitation and ability preservation. Stroke is a classic case of event impairing balance, for instance for those patients who manifest hemi-paresis which then results in minor postural control [118]. It is worth noting that more than half people surviving a stroke experience motor deficits due to reduced balance [119]. Another disease that typically impairs balance is Parkinson's. Here, patients manifest slowness of movement, rigidity of muscles, and postural instability [120], and such symptoms are destined to worsen due the progressive nature of the disease. As a result, patients experience loss of balance, leading to frequent falls and to low independence on activities of daily living [121, 122]. In general, several clinical conditions affecting balance are strictly linked to the age, thus frail elders result in a population which typically presents balance impairment. It is worth mentioning that falls are the leading cause of fatal and non-fatal injuries among seniors [123].

The study of balance impairment is therefore raising great interest also from a technological point of view, in order to provide therapists with much information. Several researches are focused on assessing balance properties for driving and forecasting the rehabilitation activities [124, 125, 126, 127, 128, 129]. In addition, from the point of view of the rehabilitation process itself, robotic devices and wearables based solutions are proposed [130, 131, 132, 133].

The most ubiquitous of technology for balance study consists on the development of platforms delivering activities, coupled with some sensor or controller for user interaction and data collection. The most used devices include Microsoft Kinect [134], Nintendo Wii balance board [135], and robotic systems [136], and are involved in studies for a wide spectrum of diseases, including stroke [118], Parkinson's [137], and frail elders [138].

CHAPTER 3

Image processing techniques for medical diagnosis

In the present chapter, the theoretic fundamentals for the developed methods are introduced. In particular, details on graphs and statistical techniques are revised. Then, the methods used in the study related to raters-agreement problem are briefly described. In the end, a brief recap of the classical image processing techniques used for the unsupervised evaluation of segmentation results concludes the chapter.

3.1 Graph notation

A graph is a pair $G = (I, E)$ where $I = v_1, \ldots, v_n$ is the set of vertices[1] and $E \subseteq \{(v_i, v_j) \in I \times I \ \mid v_i \neq v_j\}$ is the set of edges.

A graph is said *weighted* when either the vertices or the edges are labeled with weights. In the former case, the graph is called a vertex-weighted graph; in the latter case, the graph is called an edge-weighted graph. The weights can be either positive or negative and can be assigned by different functions and with different criteria.

Two vertices v_i and v_j are *adjacent* or *neighbors* if an edge $e_{ij} = (v_i, v_j)$ connecting v_i with v_j exists. A graph G is indirect (or not oriented) if the pairs of the edges are not ordered, i.e., $e_{ij} = e_{ji}$. The set of points which are neighbors of a vertex v (that is, the neighborhood of v) is denoted by G_v. A subset $C \subseteq I$ is a *clique* if every pair of distinct vertices in C are neighbors.

Let $\pi(v_i, v_j) = \langle v_1, \ldots, v_k \rangle$ be an ordered sequence of vertices. Then π is a *path* from v_i to v_j if for any $i \in [2, \ldots, k]$, v_i is adjacent to v_{i-1}. If the extremities of the path coincide, π is a *cycle*. The graph G is said *connected* if any two vertices of G are linked via a path.

---

[1] in Section 5.1, vertices are indicated as letters without ordering subscript for ease of notation. Conversely, here the subscript enables a better readability and thus it is used.

Let $G = (I, E)$ and $G' = (I', E')$ be two graphs; if $I' \subseteq I$ and $E' \subseteq E$, then G' is a *subgraph* of G, and is noted $G' \subseteq G$. If a subgraph G' is a connected graph without cycle, then G' is a *tree*. A *spanning tree* is a tree containing all vertices of the graph. A Minimum Spanning Tree (MST) is a spanning tree for which the sum of the edges is minimal, i.e.,

$$c(T) = \min \sum_{l \in path} w_l, \tag{3.1}$$

where T is the tree, $c(T)$ is the *cost* associated to T, i.e., the sum of the weights along T, and l can be either weighted edges or vertices, w_l represents the weight assigned to l.

A Shortest Path Tree (SPT) *rooted* at a vertex r defines a tree composed by the union of the paths between r and each of the other vertices in G such that

$$c(r, v) = \min c_G(r, v), \tag{3.2}$$

where $c_G(r, v)$ is the cost of the path, i.e., the sum of the weights of the edges belonging to the path.

3.2 Markov random fields

A brief introduction to Markov random field (MRF) is here provided, following the dissertation in [139].

Image labeling is a general framework in image processing, consisting of associating each pixel in the image with a label from a finite set. The meaning of such label depends on the problem under analysis, and could either be semantic, as in classification tasks, or numerical, as in segmentation tasks, where it represents the belonging region.

There are several methods for choosing the label. A widely used one consists of building a probabilistic image model and select the most likely labeling. In a probabilistic framework, pixels similarities are expressed by Markov random field. A random field is a generalization of a random process, where a collection of random variables is defined in a multi-dimensional domain. Hence, a label map can just be thought as a realization of a random field of discrete random variables.

Before tackling into details the Markovian modeling theory, it is worth spending some words on the definition of the maximum a posteriori (MAP) estimate, both in the case of independent and dependent pixels of the image.

Let $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_M\}$ be the set of feasible labels. By evaluating each element of the image in an independent way, the MAP classifier assigns to a pixel x^* the label λ_i^* that verifies the condition:

$$\lambda_i^* = \arg\max_{\lambda_i \in \Lambda} p(\lambda_i | x_i = x^*). \tag{3.3}$$

By means of the Bayes theorem, the posterior probability can be written as:

$$p(\lambda_i | x_i = x^*) = \frac{p_{x_i|\lambda_i}(x^*) p(\lambda_i)}{p_{x_i}(x^*)} \tag{3.4}$$

where $p_{x_i|\lambda_i}(x)$ is the so-called likelihood function and $p(\lambda_i)$ is the prior probability of the label λ_i.

In the case that labels are equiprobable, the MAP estimate results in being the maximum-likelihood estimate (notice that the term $p_{x_i}(x^*)$ does not depend on the label).

In this situation, the label is uniquely dependent on pixel intensity, while, for instance, its spatial position is neglected. Of course, ignoring this kind of information may lead to errors.

In a more general case, let $X^* \in \mathbb{R}^N$ be the vector of all the N observations and let $L \in \Lambda^N$ the random vector of the N labels composing the output map. The MAP criterion is applied to find the best configuration of labels L^* that maximizes the posterior probability function $p(L|X = X^*)$, i.e.:

$$L^* = \arg\max_{L \in \lambda^N} p(L | X = X^*). \tag{3.5}$$

The major drawback of the present approach is the total number of different configurations that should be taken into account, making it computationally demanding. By adopting a Markovian model, it is possible to marginalize the posterior probability function $p(L|X)$

in such a way that the maximization problem is dramatically simplified.

A common way to define MRFs related to image models is to refer to a graph $G = (I, E)$ where $I = v_1, \ldots, v_n$ is the set of vertices and $E \subseteq \{(v_i, v_j) \in I \times I \ \ | \ v_i \neq v_j\}$ is the set of edges.

Each vertex of the graph is assigned a label λ from a finite set of labels Λ. Such an assignment is called a *configuration*, is denoted ω, and has some probability $p(\omega)$ and belonging to the set of all possible configurations Ω. The restriction to a subset $T \subseteq I$ is denoted by ω_T and $\omega_v \in \Lambda$ denotes the label given to the vertex v.

Then, X is a Markov random field with respect to G if

- $p(X = \omega) > 0 \ \forall \ \omega \in \Omega$;
- $p(X_v = \omega_v | X_r = \omega_r, v \neq r) = p(X_v = \omega_v | X_r = \omega_r, r \in G_v) \ \forall \ v \in I$ and $\forall \ \omega \in \Omega$.

Such condition implies that, for the choice of the optimal label, the totality of the elements carries as much information as the subset of elements in the neighborhood of the considered node.

In addition, modeling the spatial-contextual correlation by means of a MRF allows for moving from the global joint maximization of the posterior function to a neighbor-local formulation of the problem. This is a consequence of the Hammersley-Clifford theorem [139]. In particular, such a theorem states that there exists an equivalency between $p(\omega|X)$ and a *Gibbs distribution*, i.e.,

$$p(\omega|X) = \frac{1}{Z} \exp\Big(- U(\omega|X)\Big), \tag{3.6}$$

where Z is a normalization term and U represents the energy associated to a given configuration.

In particular, $U(\omega)$ is expressed as:

$$U(\omega) = \sum_{S} V_S(\omega), \tag{3.7}$$

with S being one of all possible subsets of elements of the domain. The term $V_S(\omega)$ is a potential function that, given a configuration ω, assigns a value to each subset S.

The MAP criterion can be reformulated as:

$$\begin{aligned} \omega^* &= \arg\max_{\omega\in\Omega} p(\omega|X) \\ &= \arg\min_{\omega\in\Omega} U(\omega|X) \\ &= \arg\min_{\omega\in\Omega} \sum_S V_S(\omega|X). \end{aligned} \tag{3.8}$$

In this way, it is possible to explicit the contribution of the dependency among neighbor elements in the choice of the optimal configuration ω^*. In addition, reasoning in terms of potential functions is usually much easier than in terms of local characteristics and spatial correlation of the discrete random variables [139].

The energy function can be separated into two components, i.e., a unary term and a pairwise term, taking into account pixel-wise and spatial-contextual information respectively:

$$U(\omega|X) = \sum_{S-unary} V_S(\omega|X) + \sum_{S-pairwise} V_S(\omega|X) \tag{3.9}$$

Depending on the chosen potential function, different kinds of pixels correlation can be modeled. One of the most used potential functions is the Potts model. By integrating the Potts model inside the energy function $U(\omega|X)$, the pairwise term is simply related to the number of labels in the neighborhood of the image pixels. Therefore, the energy becomes:

$$U(\omega|X) = \sum_i -\ln \hat{p}(x_i|\lambda_i) + \beta \sum_{j\sim i} \Big(1 - \delta(\lambda_i, \lambda_j)\Big) \tag{3.10}$$

with $\delta(\lambda_i, \lambda_j)$ being the Kronecker function.

The minimization of the energy function can be addressed via multiple strategies, ranging from deterministic methods like the iterated conditional mode (ICM) [140] strategy, to stochastic methods like simulated annealing (SA) [139]. The former are generally faster but less accurate, while the latter are generally more effective, but they do require long computational times. In the last decade, the minimization strategy based on graph cuts has gained lot of attention because it is able to combine effectiveness and computational efficiency.

In the present case, the minimization of the energy function is a complex task, but it can be proved that the presented energy function is graph-representable [141], i.e., its minimization coincides with the identification the max-flow / minimum-cut of a graph. However, the computation of the minimum-cut for a multi-label MRF model is computationally demanding problem. Hence, for enabling a computationally feasible solution, iterative procedures of sub-optimal cuts have been designed. One of the most used ones is the $\alpha - \beta$ *swap* algorithm [142], which is the one that will be employed in the experimental phase.

3.3 Parameter estimation

This section is aimed at presenting the different model estimation strategies adopted in the method developed for the present research. The estimation strategies are here described considering a generic random variable Z and a set $\mathcal{Z} = \{z_i\}_{i=1}^{q}$ of independent and identically distributed (i.i.d.) samples drawn from Z.

Different parametric models may be embodied in the technique based on prior information on the characteristic of the considered image data. In general, any parametric model, for which a corresponding parameter estimation approach is available, can be integrated in the proposed approach. Among the available possibilities, two specific models can be considered, which encompass a significant number of image processing applications.

First, when aiming at isolating homogeneous and low-granularity regions, the case of a Gaussian model can be taken into consideration. Such a model is known to be well suited for modeling, for example, images collected by passive cameras – for which the additive Gaussian noise is usually an appropriate model – as well as MRI instruments [143, 144]. In this case, the estimation of the underlying parameters is addressed, in a maximum-likelihood fashion, via the sample mean and the sample variance.

Second, the case of images affected by a multiplicative noise-like component is addressed. A variety of parametric models have been introduced in the literature, including Gamma, Weibull, log-normal, K, symmetric α-stable, $\mathcal{G}_0$, generalized Gaussian-Rayleigh, Fisher, and generalized Gamma distributions [145, 146, 147, 148, 149].

It is worth noting that the models characterized by more than two parameters (e.g., Fisher, K, generalized Gamma, etc.) have been mostly designed for modeling non-homogeneous regions, where texture or granularity phenomena are highly visible. Indeed, leveraging once more on the homogeneity and low-granularity of the regions identified by the novel proposed methodology 5.1, the Gamma distribution has been considered here. Such a distribution is a well-known model for the statistics of homogeneous non-textured regions of image data affected by multiplicative speckle, as for example in radar imagery [150], sonar systems [151], and ultrasound imagery [152], and whose PDF is positive-valued ($z \in \mathbb{R}^+$):

$$p_Z(z) = \frac{1}{\Gamma(L)} \left(\frac{L}{u}\right)^L z^{L-1} \exp\left(-\frac{Lz}{u}\right), \tag{3.11}$$

where u is the mean of the distribution and L is its shape parameter $(L, u > 0)$.

In this case, parameter estimation can be performed via the Method of Logarithmic Cumulants (MoLC) [153]. MoLC estimation follows a principle similar to the well-known Method of Moments (MoM), which makes use of the Laplace transform (in terms of moment generating function) to define relations between the moments and the parameters of the underlying distribution. In the case of MoLC, the Mellin transform is used to relate the parameters of the distribution to its logarithmic cumulants. In the application to several of the aforementioned distributions, MoLC has led to advantages over MoM in terms of lower estimation variance and over maximum likelihood in terms of easier analytical or numerical solution [154].

In particular, if Z is a positive-valued random variable,the first- and second-order log-cumulants κ_1 and κ_2 can be proven equivalent to the logarithmic mean and variance of the distribution of Z:

$$\begin{aligned} \kappa_1 &= \mathbb{E}\{\ln Z\}, \\ \kappa_2 &= \mathbb{E}\{[\ln Z - \kappa_1]^2\}. \end{aligned} \tag{3.12}$$

Leveraging on the use of the Mellin transform and on the definition of the Gamma distribution (see Eq. 3.11), these logarithmic cumulants can be related to its parameters $[u, L]$ through the following MoLC equations [153]:

$$\begin{cases} \kappa_1 = \Psi(0, L) - \ln u - \ln L, \\ \kappa_2 = \Psi(1, L), \end{cases} \tag{3.13}$$

where $\Psi(0, L)$ is the digamma function (i.e., the logarithmic derivative of the Gamma function) and $\Psi(1, \cdot)$ is the first order polygamma function (i.e., the derivative of $\Psi(0, L)$) [155].

Given the i.i.d. samples in $\mathcal{Z}$, the sample estimates of κ_1 and κ_2 can be obtained as:

$$\begin{aligned} \hat{\kappa}_1 &= \frac{1}{q} \sum_{i=1}^{q} \ln z_i, \\ \hat{\kappa}_2 &= \frac{1}{q} \sum_{i=1}^{q} [\ln z_i - \hat{\kappa}_1]^2. \end{aligned} \tag{3.14}$$

Then, substituting the sample estimates into Eq. 3.13, it is possible to write the MoLC equations for the Gamma distribution relating the MoLC estimated parameters $\hat{u}$ and $\hat{L}$ to $\hat{\kappa}_1$ and $\hat{\kappa}_2$:

$$\begin{cases} \hat{\kappa}_1 = \Psi(0, \hat{L}) - \ln \hat{u} - \ln \hat{L}, \\ \hat{\kappa}_2 = \Psi(1, \hat{L}), \end{cases} \tag{3.15}$$

$\hat{L}$ is first obtained by numerically solving the second equation via the Newton-Raphson method [156], and then $\hat{u}$ is retrieved from the first equation by substituting $\hat{L}$. The former solution step is numerically simple thanks to the strict monotonicity of $\Psi(1, \cdot)$.

3.4 Superpixels

A group of pixels sharing some characteristics is commonly defined Superpixel. Their use in the contexts of Image Processing and Computer Vision is growing, as they enable a comprehensive understanding of local properties of the image, without basing on punctual information.

Among the methods for superpixels segmentation, two have been considered in the present work. They are described with regards to their choice for being the supporting algorithm in the annotation process. Details on their technical specifications are referred to the corresponding papers.

Linear spectral clustering: the linear spectral clustering (LSC) algorithm [157] is a superpixel segmentation method granting high boundary adherence and visual compactness. The key point for its selection is that it preserves global image structures through efficient local operations. Also, form the implementation point of view, it possesses linear computational complexity and high memory efficiency, thus not demanding too much computational efforts for a pre-processing phase.

Real-time coarse-to-fine topologically preserving segmentation: the algorithm real-time coarse-to-fine topologically preserving segmentation (TPS) [158] presents as major feature speed and accuracy. The superpixel segmentation problem is formulated in terms of boundary and topology preserving Markov random field (MRF). Analogously to the linear computational complexity and high memory efficiency for LSC method, the processing speed is a feature which makes TPS algorithm a good solution to be deployed in the data preparation phase.

3.5 Mathematical morphology

With regards to the pipeline for unsupervised evaluation of segmentation results, some operation referring to mathematical morphology are here introduced. In particular, this Section is about both binary and gray-levels dilation and erosion. As these techniques are well-known and widely covered in the literature [159], their description will be brief and minimal.

In the original paper about mathematical morphology it is stated "*the language of mathematical morphology is that of set theory. Sets in mathematical morphology represent the shapes which are manifested on binary or gray tone images*" [160]. Starting from this assumption, dilation and erosion can be defined as morphological transformations combining two sets using respectively vector addition and subtraction of set elements, i.e., the image and the structuring element. The structuring element is a binary window driving the processing. During the operational phase it is centered in each point of the image. In particular, let B be a structuring element and x be a point of the image, then the structuring element centered in x is called B_x.

The binary dilation of image A with structuring element B is defined as:

$$A \oplus B = \bigcup_{x \in A} B_x. \tag{3.16}$$

From the operational point of view, such an operator adds new pixels to the object boundaries. The dilation procedure in practice is the following:

- the central point $*$ of the structuring element is overlapped over each *true* element of the image;
- if the central point is *true* the whole structuring element is OR-ed (*true* wins) with the image;
- if the central point is $false$ the central point is zero-ed and the other points are OR-ed.

Similarly, the binary erosion of image A with structuring element B is defined as:

$$A \ominus B = \bigcup_{x \in A} B_x. \tag{3.17}$$

From the operational point of view, such an operator removes new pixels to the object boundaries. The erosion procedure in practice is the following:

- all the pixels in the image must be processed. The central point of the structuring element is applied over all the pixels of the image, from left to right, top to bottom;
- if all the *true* values of the structuring element matches the image pixels, a *true* is set where the central point is;

- otherwise a $false$ is set, eventually overwriting an existing $true$ (erosion).

When considering gray-levels dilation and erosion, the output pixel is computed as the maximum or the minimum gray level respectively, among all the pixels in the neighborhood of the point where the structuring element is centered. Therefore, such operations result in the extension of the respectively bright and dark areas of the image.

CHAPTER 4

ReMoVES and data analysis techniques

This chapter is devoted to the description to the technologies deployed for tele-rehabilitation support, including both the technical solution and the analytic technologies involved.

4.1 An IoT solution for assessment and monitoring

The present section goes over the description of the Remote Monitoring Validation Engineering System (ReMoVES) system, following the dissertation in [138].

ReMoVES was developed by the Department of Electrical, Electronics, and Telecommunication Engineering and Naval Architecture (DITEN) of Università degli Studi di Genova [161]. The proposed IoT system provides a personalized rehabilitation program that can be performed at home by the patient, while the therapist can track training performance and effectiveness from any Internet-connected device. By developing game-based rehabilitation tools that are tailored to the therapy goals of different patient categories, multidimensional rehabilitation teams can be provided with more meaningful performance data. Among others, the monitoring of eventual compensation movements allows for the evaluation of whether an exercise is correctly performed.

Several IoT architectures for telemedicine systems and e-health were proposed in the literature [162, 163, 164], but the most compliant with ReMoVES is the one composed of four layers shown in Figure 4.1.

These levels work closely together, and ensure the archiving, processing, monitoring, and proper evaluation of patients' rehabilitation performance. The four-layer architecture divides the connection part from the server/cloud part. It is important to define the correct role of the latter because the physical server used in ReMoVES plays a fundamental role in the correct processing and management of the entire IoT system.

A detailed description of the four layers follows referring to the used technologies and devices.

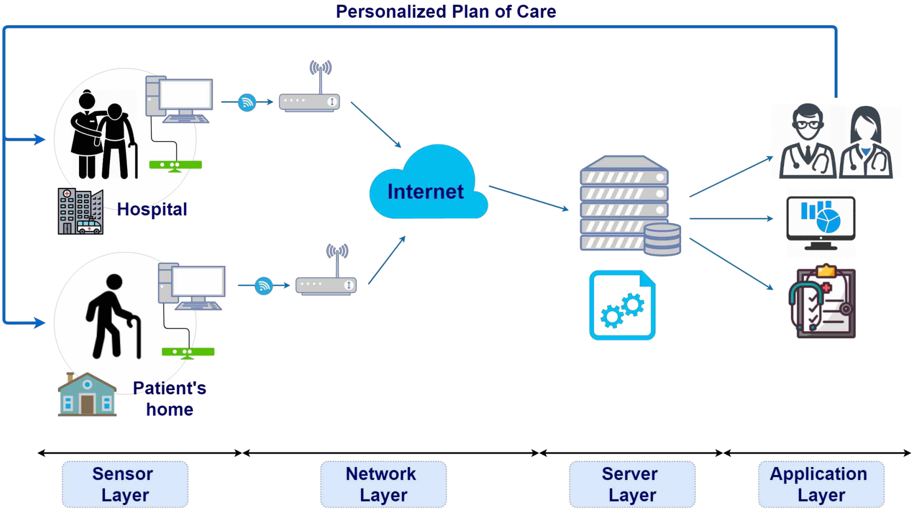


Figure 4.1: Architecture of ReMoVES . Each layer is depicted in the corresponding position.

Sensor Layer

The bottom layer is the sensor or perception layer and consists of the patient client. It deals with the management of so-called "things" (i.e., sensors connected to the system). ReMoVES employs off-the-shelf devices, i.e., Microsoft Kinect V2, Leap Motion, and a touchscreen, resulting in a low-cost solution for telerehabilitation. These devices are installed and connected to a computer, and through simple body gestures or touches (in the case of touchscreen), the patient interacts with the game shown on the screen. Patient movements are recorded without requiring the intrusive use of video cameras, which require specialized methods for tracking the user's body, and are heavy and error-prone. After the patient finishes the game session, raw information is generated from tracked data and sent to the upper level. A brief description of the included sensors in the platform is provided. The real-scenario application here refers to full-body activity; so, exergames delivered via the Kinect sensors are described, furnishing particular details about the game used for the performance assessment of frail elderly people.

Microsoft Kinect V2: A motion-sensing input device based on a time-of-flight camera to build a depth map of the environment. It can simultaneously track in 3D up to 25 fundamental joints (Figure 4.2) of the framed human body. It offers a wide field of view (70° × 60°) and recognition up to 4.5 m from the device [165]. Data from the tracked user's body are recorded at a frequency of 10 Hz. Several studies demonstrated that the Microsoft Kinect V2 can validly obtain spatiotemporal parameters [166, 167]. Microsoft Kinect is also a satisfactory tool for rehabilitation due to its low cost and adequate spatial accuracy (with an order of magnitude of centimeters) [168].

Exergames based on Microsoft Kinect have a significant field of application in assistive technologies for the elderly, such as in reducing fall risk, improving physical performance, and reversing the deterioration process in frail and prefrail elderly persons [169].

Leap Motion: Explicitly aimed at the recognition of hand gestures, it calculates the position of the fingertips and the orientation of the hand. Its deployment in ReMoVES is devoted to hand-district rehabilitation exergames.

Touchscreen: Required for interacting with the subset of exergames for cognitive assessment. The monitor is positioned on a table with an angle to the plane of a few degrees. Cognitive exergames in the ReMoVES platform are a digital reinterpretation of some gold-standard tests administered on paper to patients. Interaction through a touchscreen monitor allows for complete data collection, also helping the administrator avoid taking notes during patient activity. Examples of collected auxiliary data are interaction speed and methods or strategies used by the patient to complete the test.

Exergames: The digital games were developed from scratch for this research. They encourage the patient to autonomously carry out functional exercises along with traditional motion rehabilitation. The creation of these activities involved different processes, technologies, and specialists. It is necessary to pay particular atten-

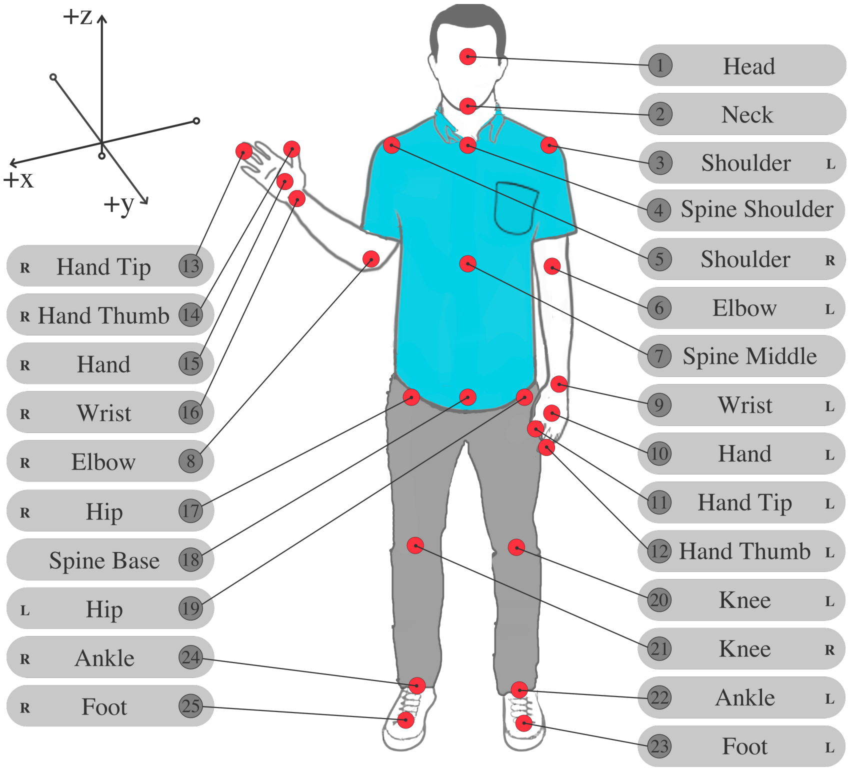


Figure 4.2: Skeleton-joint locations and names as captured by Microsoft Kinect sensor. Skeleton composed of 3D coordinates for each of its 25 joints.

tion to the specifications provided by physiotherapists and physiatrists, who share their skills to define the requirements and parameters of the game. The present exergames are considered to be assessment and rehabilitation activities, delivering task-oriented training by requiring the patients to fulfill consecutive and repetitive tasks. They foster mild-intensity activity, which promotes active aging for frail elderly individuals, and allows for the preservation or re-acquisition of functional skills that are involved in real-life activities. To design an enjoyable and safe gaming experience for elderly users, several age-related requirements needed to be considered [170]: (i) The target audience's lack of previous gaming experience: devices such as Microsoft Kinect enable users to control and naturally interact with exergames without the need to physically touch a game controller or object of any kind. Microsoft Kinect achieves this through a natural user interface by tracking the user's body movements. (ii) Exergames should focus on a simple interaction mechanism, while complex and decorative graphics should be kept to a minimum. (iii) Exergames should avoid frustration and foster an enjoyable player experience: when the motor skills of the user are reduced, a preventive calibration phase allows for the patient to complete the game task even with a limited ROM.

The system currently includes six different exergames for the Kinect sensor, five exergames for the Leap Motion sensor, and three tests for the touchscreen. The exergames can be modified on the basis of level parameters, duration, range of motion, speed, or others. These activities automatically adapt to the patient capabilities due to a calibration phase. The thumbnails of Kinect, Leap Motion, and touchscreen exergames are shown in Figures 4.3, 4.4, and 4.5 respectively.

For the sake of completeness, a brief description of the exergames follows.

- **Equilibrium Paint**: this game is an interactive version of the sit-to-stand exercise. The user repeatedly stands up and sits down within a predefined amount of time (30 s). The scene shows a horizontal wooden beam on which paint cans are placed. The inclination of the beam directly depends on the angle of the patient's shoulders during the STS, traced by Microsoft Kinect. When the patient does not symmetrically stand up, the paint cans fall down, causing a score penalty in the game.

- **Owl Nest**: the patient is encouraged to reach an on-screen target with an arm motion (reaching task) in order to achieve a high in-game score. Many colorful owls randomly appear in a position in the screen, and the user carries them to the nest to gain points. Then, other ones appear on the screen.

- **Shelf Cans**: introduces the patient to a virtual environment that is similar to a kitchen. With an arm movement, the patient grabs one of the colorful drink cans appearing in the middle of the screen and drags it to the corresponding shelf. This game

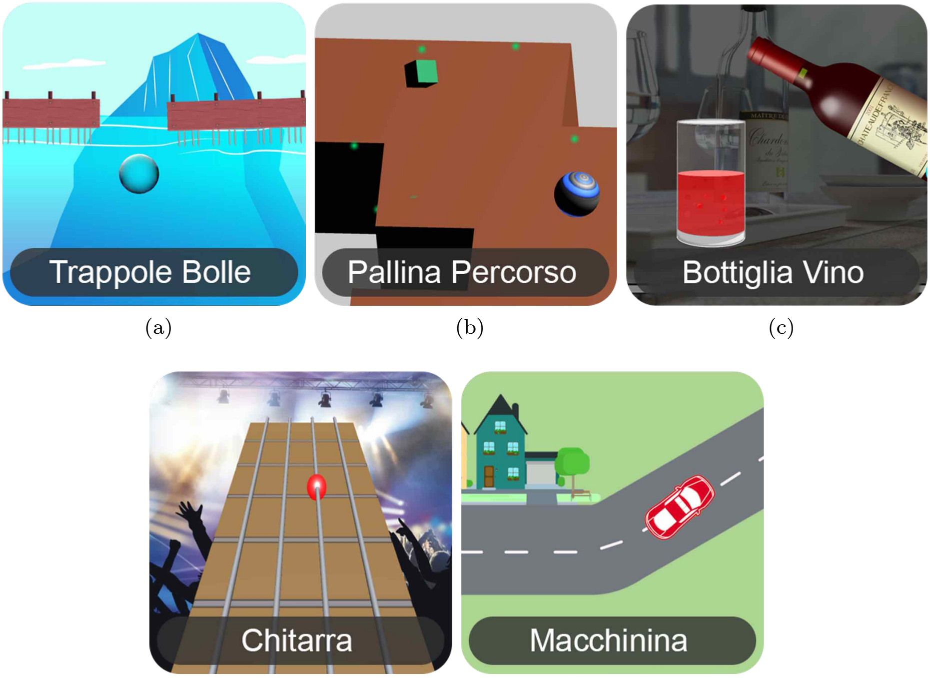
(a) (b) (c)

(d) (e) (f)

Figure 4.3: Thumbnails of Kinect exergames in current ReMoVES catalog. (**a**) Equilibrium Paint; (**b**) Owl Nest; (**c**) Shelf Cans; (**d**) Hot Air; (**e**) Push Box; (**f**) Flappy Cloud.

is appealing because it requires the user to be attentive to drop off the drink can on the correct shelf according to its color.

- **Hot Air**: this is an activity to train the patient's body balance. The user can control the direction of a hot-air balloon floating in the sky with the balance shift: in-game scores are collected when it is led towards the bonus targets.
- **Push Box**: assesses balance, where the patient must stretch forward with their arms parallel to the ground. It takes inspiration from a phase of the Berg balance test. The purpose of this exercise is to push a box into a hole a few meters in front of the box.
- **Flappy Cloud**: this is a functional exercise for the lower limbs.



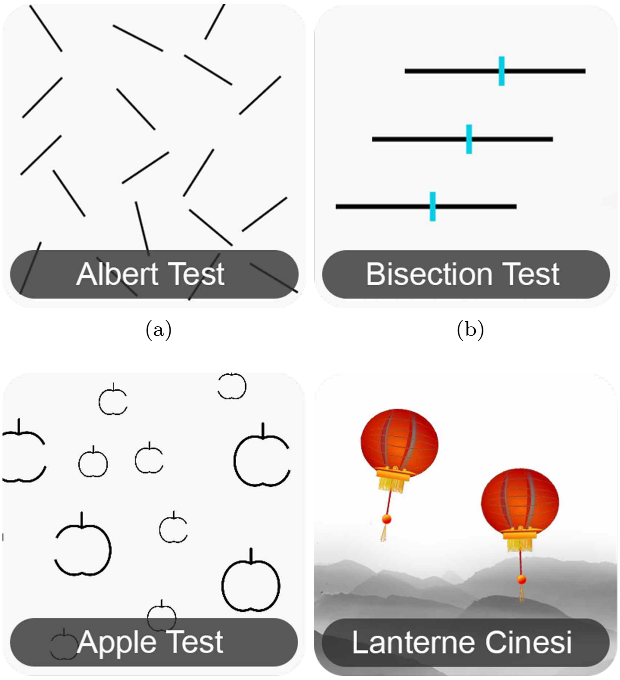
(a) (b) (c)

(d) (e)

Figure 4.4: Thumbnails of Leap Motion exergames in current ReMoVES catalog. (**a**) Floating Trap; (**b**) Endless Zig; (**c**) Wine Bottle; (**d**) Finger Tap; (**e**) City Car.

The leg abduction-adduction movement reflects the position of a cloud object in the game screen: the patient makes it move forward without hitting some obstacles.

- **Floating Trap**: the patient is led to open the hand and make a fist alternatively. This exercise requires a good level of concentration: in fact, the user moves a floating raft on the left or on the right according to the finger Flexion–Extension in order to avoid some objects in the scene.
- **Endless Zig**: the patient drives a marble along a zigzag path appearing on the screen. Going out of the boundaries causes score loss; similarly, some bonus gems appear on the path. The patient controls the marble movement with Radial–Ulnar devi-



(a) (b)

(c) (d)

Figure 4.5: Thumbnails of touchscreen exergames in current ReMoVES catalog. (**a**) Albert test; (**b**) Bisection Test; (**c**) Apple Test; (**d**) Lantern.

ation.

- **Wine Bottle**: this exercise mimics a real-world scenario, i.e., pouring liquids from a bottle. With the Pronation–Supination movement of the hand, patients should control the rotation of a bottle of wine appearing on the screen. They must fill a glass over and over again to collect as many points as possible.

- **Finger Tap**: the patients perform the finger opposition exercise, namely they are required to touch with the thumb other fingers, one finger at a time. The scene of the game represents a neck of a four-string guitar, where patients pretend to play the instrument. Some colored marbles sequentially fall off the strings. The color sequence is green, yellow, red, and blue, corresponding to fingers from left to right. The exergame self adapts to both left and right hands, by defining the correct correspondence finger opposition - color.

- **City Car**: the patient drives a car along a randomly-generated road. The user should steer in the presence of curves and crossroads with the movement of Flexion–Extension of the wrist. Penalties are introduced when the user goes off-track.
- **Albert Test**, **Bisection Test**, **Apple Test**: these are a computerized version of the well- known paper-and-pencil test for USN [171, 86].
- **Lantern**: in this exergame, Chinese lanterns appear from the lower side of the screen and rise towards the sky (upper part of the screen). The player must act as quickly as possible and touch them using on the screen. This activity is for the assessment of mental alertness and awareness. Moreover, it can have applications for the neglect assessment by comparing the performances relative to the targets shown on the right or left of the screen.

Network Layer

The role of the network layer is to establish communication between data tracked by the sensors and stored in a local PC and the remote server or cloud. In ReMoVES , data-log files in JavaScript Object Notation (JSON) format are temporarily stored in the local unit or PC installed in the patient's home or at hospital. These data are sent to the central server as soon as an Internet connection is available via Ethernet or Wi-Fi. This functionality was added in order to combat any possible connection trouble and to facilitate domestic use where a reliable Internet connection may not be available.

Server Layer

The server layer provides data elaboration and analysis via cloud or server storage. Software running on the physical ReMoVES server can manage content-independent data flow to be compliant with software reuse logic. Server software consists of a traditional Linux-Apache-MySQL-PHP (LAMP) stack, and provides data-storage solutions, data-processing methods, and a web application for clinicians to view information through dedicated graphic interfaces. The ReMoVES server has only three types of application programming interfaces (APIs) for the management of client or server data synchronization. Data communication is in secure mode based on hypertext transfer protocol secure (HTTPS). In HTTPS, the communication protocol is encrypted using transport layer security (TLS). Certificates are issued by the Let's Encrypt authority. To process acquired information, a complete post-movement reconstruction of in-game events is allowed. Additionally, this component runs the data-processing algorithms and provides the interface for displaying the results.

Database: This subsection describes the MySQL relational database used by ReMoVES . The dataset consists of a structured collection of JSON files, each of them containing an array of temporal events. In each element of the array, there are key-value pairs that provide data. Some keys are common to all exergames, such as time of recording in milliseconds (ms), in-game score, and Kinect joint position (see Figure 4.2). In addition, other keys are provided depending on the game.

Application Layer

This layer consist of the therapist client, which provides therapists, physiotherapists, and doctors with direct access to data. Specifically, the built-in algorithms provide a clear and concise report to the therapist in order to facilitate the interpretation of therapy evolution. The web interface provides a user-friendly means for the clinical staff to consult information, also displaying patient performance in graphic mode, and to assign rehabilitation therapies. The layout dynamically adapts to the size and type of device; this allows for connection even from a smartphone in the case that the therapist does not have an available computer.

Figures 4.6 and 4.7 show pictures of hardware and software architectures, respectively. The patient client is composed of a local unit with the following hardware requirements:

- processor, seventh generation Intel® Core™i5 CPU (quad-core 2.4 GHz or faster);
- memory, 4 GB RAM;
- storage, 20 GB;
- video card, DirectX11-capable from NVIDIA, AMD, or Intel with at least 1 GB VRAM; and
- dedicated USB3 port.

Microsoft Kinect or Leap Motion sensors are connected to the computer on the basis of the therapist's recovery plan; thereby, exergames are assigned to the patient. A touchscreen monitor is added instead in the case of assessment through cognitive tests. As aforementioned, an Internet connection is not mandatory for the user identification phase and to locally start the exergames, but it is necessary for data synchronization with the server. The central unit is composed of the ReMoVES server, which performs the data synchronization with the patient or client units, and stores and processes data in the MySQL database. Therapists can access from any device the web application supplied by WEB server functions.

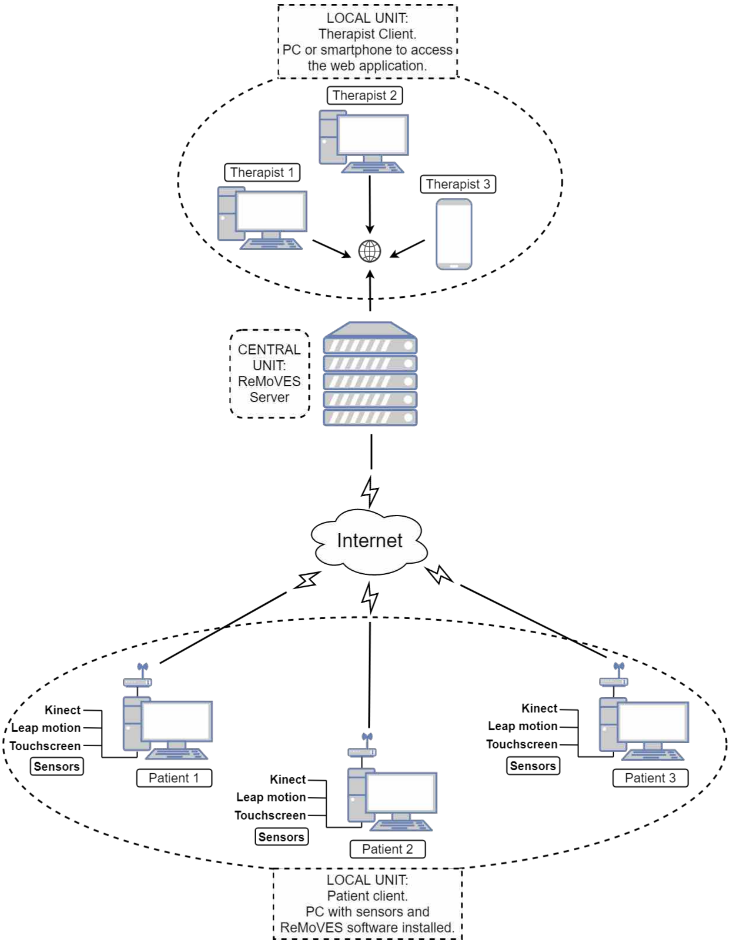


Figure 4.6: Hardware architecture of ReMoVES .

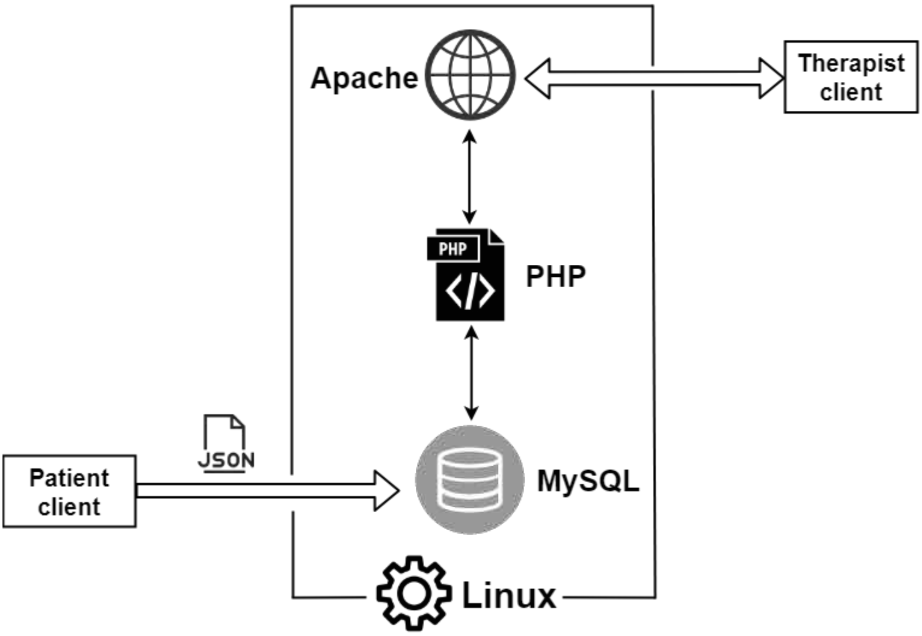


Figure 4.7: Software architecture of ReMoVES .

4.2 The project STORMS

In addition to the standard version of ReMoVES , an ad-hoc version was developed in the context of the Project Solution Towards Occupational Rehabilitation in Multiple Sclerosis (STORMS). Here, such a specific version is introduced. It includes some exergames from the standard version (i.e., Hot Air and Shelf Cans), whose description was provided in the previous Section. In addition, three new exergames was developed, with different levels corresponding to different degree of difficulty. In the end, the exergame Owl Nest was extended, so that here it includes four levels. Such novel activities are disclosed in the present Section.

The new set of exergames is mainly focused on cognitive recovery rather than motor rehabilitation. The main cognitive functions involved in games include attention, memory and executive functions. Among the sub-categories of such abilities, the games will address *working memory*, *inhibition control*, *selective attention*, *task switching and cognitive shifting*, *multitasking*, *sustained attention*, *top-down attention task* [113].

A brief description of the exergames and their levels follows.

- **Owl Nest** (STORMS): as in the standard version, the goal is to grab the owls that randomly appear on the screen with the flexion-extension of the arm and bring them into a nest placed in the middle of the screen.
  First level: no more than three owls can appear simultaneously, with no distracting elements. When the user brings an owl into the nest, another owl appears at a different point and the in-game score increases.
  Second level: some eagles appear on the screen as distracting elements. No more than five owls and three eagles can appear simultaneously. The time between the appearance of two consecutive eagles randomly ranges from 0 to 5 seconds. After 10 seconds the eagle disappears. Catching an eagle makes the game score decrease.
  Third level: the player is required to catch only the pink owls and do it as quickly as possible, since after 15 seconds the owls will disappear. Every time an owl disappears, another reap-

pears in a different location until a maximum of seven owls are simultaneously on the game scene. The score decreases either when grabbing a blue owl or when a pink owl disappears.
Fourth level: it is a combination of all the goals of the previous levels. No more than four owls and three eagles appear simultaneously. The user must bring the pink owls into the nest and avoid both eagles and blue owls. Eagles disappear after ten seconds, while owls disappear after 8 seconds.

Figure 4.8 shows some game levels

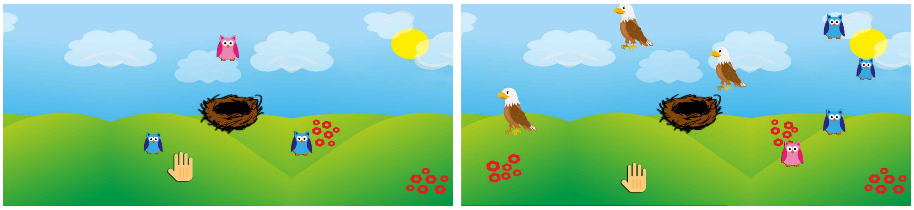

(a) Level 1 - level 3.
(b) Level 2 - level 4.

Figure 4.8: Screenshots of some levels of Owl Nest exergame.

- **Supermarket**: this exergame is set into a supermarket where the player is instructed to buy some objects. A list of items to pick up will be displayed at the start of the game. The user has between 8 and 25 seconds to memorize this list. When the game starts the patient, with the movement of the arm, will have to take the correct objects. An audio feedback is provided, with a positive or negative sound occurring in case of correct or incorrect action, respectively. In the first, third and fourth levels, once the list has been completed and the patient collected all the objects, these ones will reappear on the screen changed position. The user will then be able to collect the stored objects again.
  First level: a temporary list of three food names must be memorized in eight seconds. Then, the player has to collect the relative objects, located on two lateral shelves, into the shopping bag, in the middle of the screen. Some non-food distractors appear on the shelves. The semantic property (food or non-food)

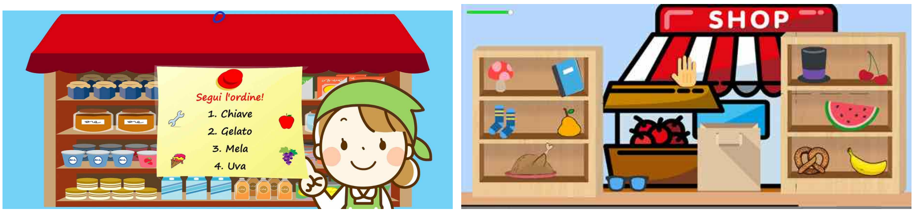
of the objects is the crucial correctness factor of the activity. The game score decreases if wrong objects are put into the bag. A visual feedback (health bar initially shading to yellow and red as mistakes are made up) is provided as well.

Second level: the player has to memorize in order four objects in ten seconds. These objects must be collected among other distracting items sliding on the conveyor belt. The game score decreases if the distractor object is taken or if the order is wrong.

Third level: the player must memorize (in twenty seconds) and follow four sequenced instructions. The sequence of the instructions must be respected. Each instruction refers to a different semantic characteristic (shape, color, or material) of the objects to be collected. Collected items will disappear, only to reappear again with the start of a new round.

Fourth level: it is like the third level, but it differs for the higher number of the objects in the scene and the number of instructions (five instead of four). In addition, the objects will reappear on the screen once collected.

In Figure 4.9 the screenshots of the four levels are provided.



(a) List of items in the Level 2.

(b) Level 1.

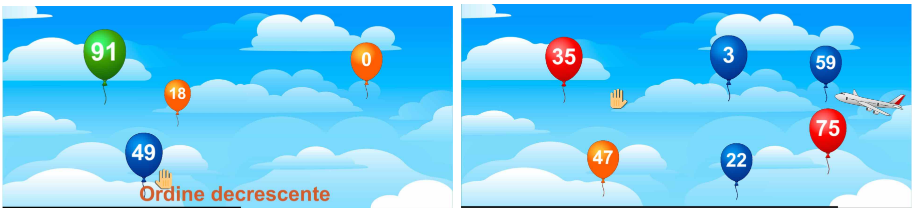
(c) Level 2.

(d) Level 3.

Figure 4.9: Some screenshots of the levels of Supermarket exergame.

- **Numbers**: the patient has to pop some numbered balloons (from 0 to 99), according to temporary instructions. Four instructions alternate according to the level difficulty: i) pop the balloons in ascending order; ii) pop the balloons in descending order; iii) pop the balloons with even numbers; iv) pop the balloons with odd numbers. The number of balloons varies depending on the difficulty level, they also have different colors and sizes to make the game more dynamic and visually more appealing. If the user takes a wrong balloon, a red mark appears at the bottom of the screen, otherwise if he catches all the correct balloons a green mark appears. In both cases, all the balloons still on the screen will be destroyed and a new round will begin with new balloons.
  First level: four balloons appear on screen. They must be popped in either ascending or descending order.
  Second level: it is as the first level, except that the patient is also required to pop either the even or the odd balloons and there are five balloons in the game scene. Two visuo-verbal stimuli are added. When the text relating to the assignment "take the odd numbers" shows, a red bird will appear for a few seconds flying from one side of the scene to the other. Conversely, a plane will appear on the screen for a few seconds when the text relating to the task "take the even numbers" appears.
  Third level: six balloons are simultaneously displayed. Once more, all the tasks can be performed, but this time the patient must remember the stimuli association previously described, because in the cases of "take the odd numbers" and "take the even numbers" no writing appears on the screen.
  Fourth level: it is structured like the third level, but also the writings "pop in ascending/descending order" will disappear after a couple of seconds.
  Fifth level: the player user has to quickly pop as many correct balloons as possible before they disappear. In fact, all the four balloons will fly off and disappear from the screen. Only the tasks about the odd and the even numbers with their relative visuo-stimuli, will be displayed.

Figure 4.10 depicts the screenshots of some levels.



(a) Level 1.

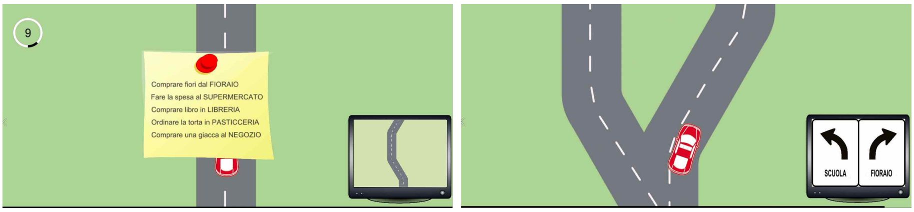
(b) Level 3.

Figure 4.10: Screenshots of the level one and level three of Numbers exergame.

- **Business By Car**: the patient drives a car along a randomly generated road. In particular, the car turns either left or right as the player moves the trunk laterally to the left or to the right respectively. The speed of the car increases progressively and returns to the initial condition as soon as the player goes out of the carriageway, introducing a penalty in the score. In this last case the car will be re-positioned on the path. At the beginning of the game, a list of places to visit appears. The patient will have to memorize this list in a time that varies between ten and twenty seconds, based on the selected level. Then, the game starts and the patient has to drive the car along the path and select the correct street at the crossroads to pass by the required places. In the end, a series of multiple choice questions will appear on the screen, either related to the list and places or about details present in the game scenes or in the buildings visited. To answer the questions the patient has to raise his arm and guide the hand that appears on the screen towards the answer button. In detail, the three levels will be described which differ according to the number of places to remember or the difficulty of the final questions.
  First level: in the easiest level, the patient must remember only four places to visit. The buildings to see are simply written on the list, with no further writings, to make the goal clear. At a crossroads, the navigator at the bottom-right will indicate the correct route to take. If the patient were to take the wrong path, a message will appear, reminding him of the correct place

to visit and to pay more attention to the next crossroads. In the question scene, there will be two questions.
Second level: in this level the patient must visit five places. This time a real list of errands to be carried out will appear at the beginning of the game. The navigator does not indicate the correct path to take and only a warning message will appear on screen if the player goes on the wrong direction. The final questions are three. *Third level:* in the last level the player must keep in mind six places. No warning message will appear if he takes the wrong way. In the final scene, the patient should answer to four questions.

Figure 4.11 shows some screenshots of the exergame.


Comprare fiori dal FIORAIO
Fare la spesa al SUPERMERCATO
Comprare libro in LIBRERIA
Ordinare la torta in PASTICCERIA
Comprare una giacca al NEGOZIO

(a) List of errands to do.

(b) The crossroad with the directions given by the navigator.

(c) The building to visit.

Figure 4.11: Screenshots of the Business By Car exergame.

4.3 Support vector machine

A detailed dissertation to SVM can be found in [172]. Here, a very short introduction is provided.

A support vector machine is a supervised learning technique from the field of machine learning applicable to both classification and regression. It is mainly used for binary classification, consisting on a non-probabilistic binary linear classifier.

The key idea of SVM is to create a plane (or hyperplane) separating the two classes considered. More specifically, the hyperplane to be defined is the one maximizing the separation between the classes. It is hence clear that SVM results in an optimization problem, in particular, in a constrained minimization problem [172].

Let C_1 (◯) and C_2 (□) be the two classes of points to be separated. Assume they are linearly separable and let them be marked with symbolic labels $y = +1$ and $y = -1$ through a linear separator π as in Figure 4.12.

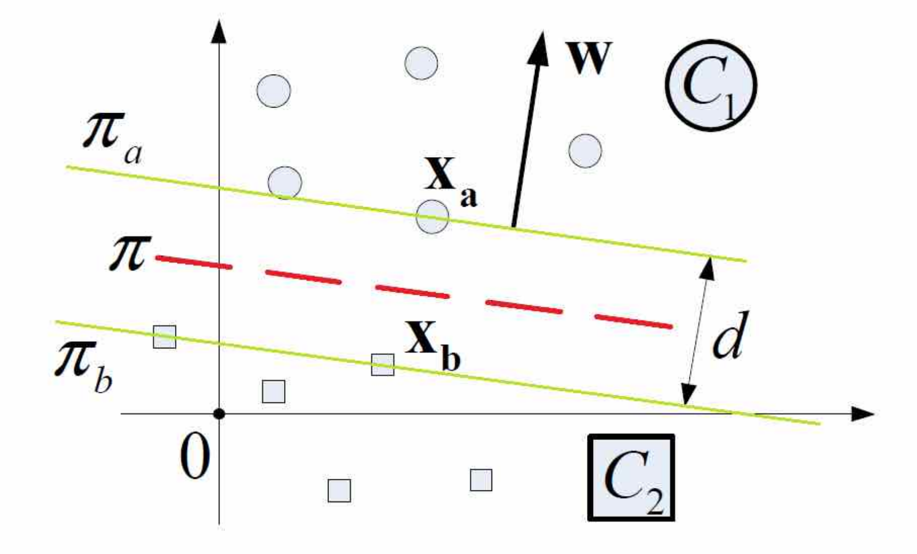


Figure 4.12: Points in C_1 and C_2 and separation region for a given $\mathbf{w}$.

Among the several planes dividing such classes, the aim of SVM is to find the one maximizing the distance d between two support planes π_a and π_b. Let $y = \mathbf{w}^t\mathbf{x} + b$ be a generic plane function, where $\mathbf{w}$ is a vector perpendicular to the plane and b indicates the translation of the plane in respect to the origin. The two support planes π_a and π_b are defined with the equations:

$$\begin{cases} \pi_a : \mathbf{w}^t\mathbf{x}^+ + b = +1 & (\mathbf{x}^+ \in C1) \\ \pi_b : \mathbf{w}^t\mathbf{x}^- + b = -1 & (\mathbf{x}^- \in C2) \end{cases} .$$

Considering then $y^+ = +1$ the label of (◯) and $y^- = -1$ the label of (☐):

$$\begin{cases} \mathbf{w}^t\mathbf{x}^+ + b \geq 1 \\ \mathbf{w}^t\mathbf{x}^- + b \leq -1 \end{cases} . \tag{4.1}$$

Assuming that the domains of points is of the form $D_n = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, a general expression for the constraint is

$$y_i(\mathbf{w}^t\mathbf{x}_i + b) \geq 1 \ \forall i \in \{1 \ldots n\}$$

Recall that the distance between π_a and π_b is $\frac{2}{\|\mathbf{w}\|}$; hence, the goal is to find the parameters $\mathbf{w}$ and b such that this distance is maximal. This is the constrained optimization problem, which leads to solve the equation

$$\mathbf{w} = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i, \tag{4.2}$$

where the α's are obtained by maximizing

$$\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j \tag{4.3}$$

subject to $\alpha_i \geq 0$ and $\sum_{i=1}^{n} \alpha_i y_i = 0$.

This is a quadratic programming problem, thus a global maximum can always be found.

To conclude, b can be determined by the equation

$$\mathbf{b} = y_k - \sum_{i=1}^{n} \alpha_i y_i \langle \mathbf{x}_i, \mathbf{x}_k \rangle. \tag{4.4}$$

Therefore, new classification tasks can be performed, via the scheme:

$$\mathbf{x} \longrightarrow \mathbf{w}^t\mathbf{x} + \mathbf{b} \longrightarrow \begin{cases} \geq 1 & \mathbf{x} \in C_1 \\ \leq -1 & \mathbf{x} \in C_2 \end{cases} .$$

In a more general situation, classes are not separable by a hyperplane. However, linear classifier can still be used by considering some data points like errors due to noise.

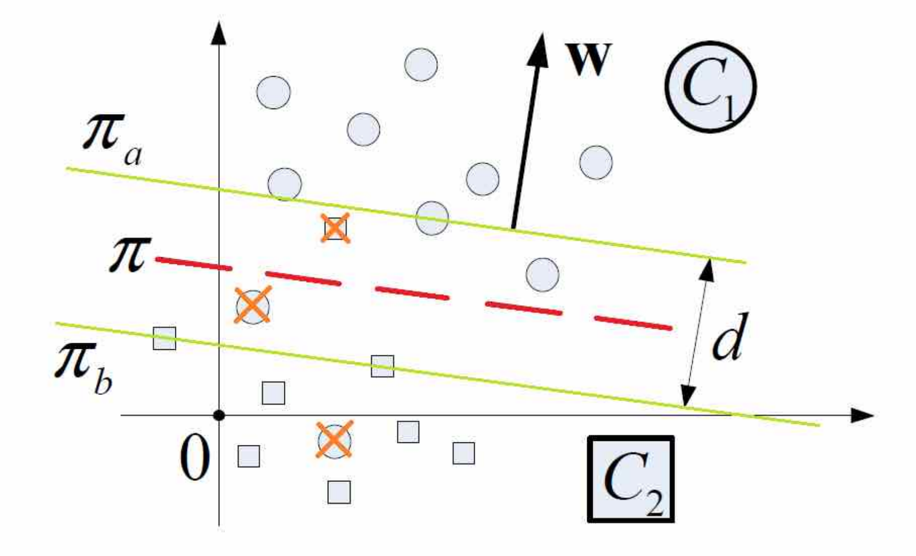


Figure 4.13: The orange crosses are wrong points that violate the constraints.

Here, some violation of constraints should be allowed. This can be done by introducing the *slack* variables ξ_i.

In the present case the problem of minimization becomes:

$$\min_{\mathbf{w},b,\boldsymbol{\xi}} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i \qquad \begin{cases} y_i(\mathbf{w}^t\mathbf{x}_i + b) - 1 + \xi_i \geq 0 \\ \xi_i \geq 0 \end{cases} \tag{4.5}$$

with C prefixed > 0.

Without going into details the solution of this minimization problem is similar to the previous case, with the difference that the parameter b depends on both $\boldsymbol{\alpha}$ and $\boldsymbol{\xi}$.

A further way to deal with data which are not linearly separable is to perform the classification via a non-linear function. This implies passing from a space L of vectors $\mathbf{x}_i$ to a space H in which points are linearly separable $L \ni \mathbf{x} \mapsto \phi(\mathbf{x}) \in H$.

Then a linear separation can be done with points $\phi(\mathbf{x}_i)$. In general, the H space has greater dimension than L space. An example of such a situation is provided in Figure 4.14, where where:

$$\underbrace{\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix};}_{L=\mathbb{R}^2} \qquad \underbrace{\Phi(\mathbf{x}) = \begin{bmatrix} \mathbf{z}_1 = \mathbf{x}_1^2 \\ \mathbf{z}_2 = \sqrt{2}\mathbf{x}_1\mathbf{x}_2 \\ \mathbf{z}_3 = \mathbf{x}_2^2 \end{bmatrix}}_{H=\mathbb{R}^3}$$

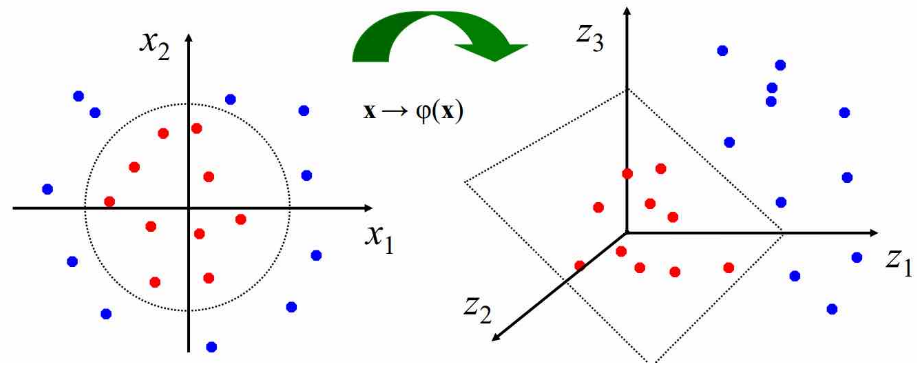


Figure 4.14: Example of space mapping for enabling linearity.

Considering the dot product $\mathbf{x}_i^t\mathbf{x}_j$, the application the non-linear function increases the computational cost. However, data points are involved in the minimization problem only for dot product. Hence, to avoid computational issues one can leverage on the the so-called *Kernel Trick*, i.e., use a kernel function K so that:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^t\phi(\mathbf{x}_j). \tag{4.6}$$

The most commonly used kernel functions are:

- Polynomials:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\langle\mathbf{x}_i, \mathbf{x}_j\rangle + 1)^d \qquad d = 1, \ldots.$$

- Radial Basis Functions (RBF):

$$K(\mathbf{x}_i, \mathbf{x}_j) = exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \qquad \text{“}\mathit{Gaussian}\text{”}.$$

4.4 Long short term memory recurrent neural network

The large dimensionality of data is one of the most significant challenges in pattern recognition. Indeed, as the number of extracted features in the input increases, so does the number of training samples required for the training system. That is, as the number of features increases, the system prediction performance may decrease. Such a problem is known as the *curse of dimensionality*, and led to the approach typically used for pattern recognition, i.e., dividing the system into two modules: a feature extraction to reduce dimensionality, followed by the classifier.

As major drawback of such an approach, accuracy and precision of the predictions are strongly influenced by the ability of the person who creates the feature extraction module, which also requires a considerable knowledge of the application domain.

In such a context deep learning techniques allow for overcoming this problem through methods and algorithms of data representation working in a hierarchical manner on several levels of abstraction. The feature representation is learned directly from data, through a series of non-linear transformation. This allows for avoiding manual implementation, which is the key aspect of deep learning approaches.

Deep learning techniques deal with neural network (NN), which, in general, are divided into convolutional neural networks and recurrent neural networks. As for the former, convolutional neural networks typically deal with image-data and perform convolutionary operations. They extract both general and particular features in images, depending on their depth, thus are used in image recognition and segmentation, style transfer, etc. Conversely, recurrent neural networks perform the same task for each sequence element, with the output depending on the previous calculations. They can use information in arbitrarily long sequences, but in practice they are typically limited to looking back only a few steps.

A particular RNN architecture is introduced in this section. Usually, in neural networks each input and output is assumed to be independent from the others. Nevertheless, such hypothesis is sometimes not suitable for many types of issues, for instance in the application of speech recognition where knowing the preceding word is undoubtedly essential. The name *recurrent* indeed is suggesting the nature of the computations RNNs perform.

In such a framework, a widely used network is the long-short term memory network [67].

LSTM networks were introduced as a solution to the vanishing/exploding gradient problem. In RNNs, the calculation and back-propagation of gradients may lead to issues, as they tend to either increase or decrease at each time instant and therefore, after a certain number of iterations, they diverge to infinity or they converge to zero.

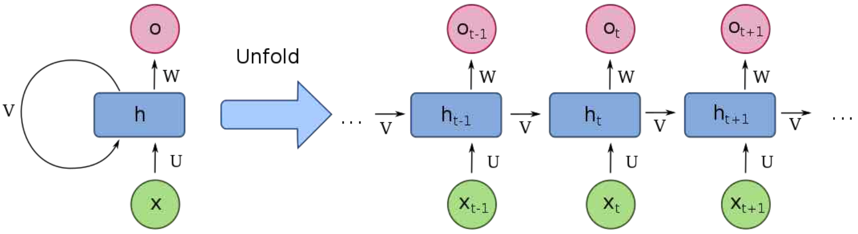


Figure 4.15: Scheme of RNN

In Figure 4.15, showing a general RNN scheme, x_t is the network input at the instant t and h_t is the state vector, which can be considered a sort of *memory* of the system as it contains information of all the previous elements of the input sequence. The vector h_t is calculated starting from the current input and from the state vector in the previous instant, through the matrices U and V as follow:

$$h_t = f(Ux_t + Vh_{t-1}), \tag{4.7}$$

where the function f is a non-linear function. Finally, o_t is the output of the network at the instant t, calculated by the parameter W.

The LSTM networks, instead, are equipped with special hidden units, called *memory cells*, whose task is to remember the previous inputs for a long time and preserve the error that can be back propagated through layers.

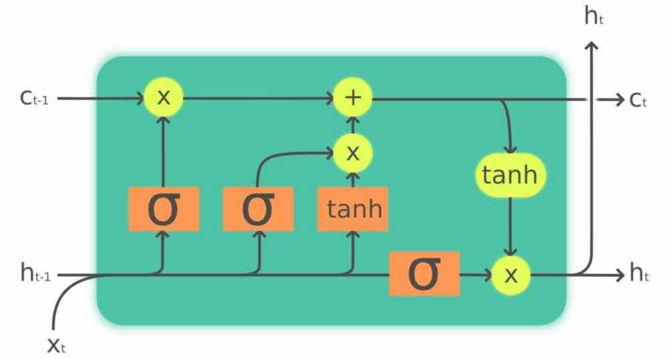


Figure 4.16: Scheme of LSTM.

A common architecture of LSTM [173, 174], showed in figure 4.16, is composed of a cell state and three structures called gates which remove or add information to the cell state itself. The gates are composed out of a sigmoid neural net layer (σ) and a point-wise multiplication operation. The sigmoid layer outputs numbers between zero and one, describing how much of each component should be let through.

First, a sigmoid layer called the *forget gate layer*, decide what information should be thrown away or kept from the cell state. Information from the previous hidden state h_{t-1} and information from the current input x_t is passed through the sigmoid function. Values come out between 0 and 1. Getting closer to 0 means *forget*, and getting closer to 1 means *keep*. The output of this first step is

$$f_t = \sigma(W_f[h_{t-1}, x_t]) + b_f. \tag{4.8}$$

Then, a *input gate layer* update the cell state. First, the previous hidden state and current input are passed into this sigmoid function. Next, a hyperbolic tangent layer creates a vector of new candidate values, $\tilde{C}_t$, that could be added to the state. The function tanh helps to squish values between -1 and 1 to regulate the network. Then, the output is multiplied with the sigmoid output i_t. The latter decides which information is important to keep from the tanh output.

$$i_t = \sigma(W_i[h_{t-1}, x_t]) + b_i. \tag{4.9}$$

$$\tilde{C}_t = tanh(W_C[h_{t-1}, x_t]) + b_C. \tag{4.10}$$

Now, the new cell state can be calculated. The old cell state gets point-wise multiplied by the forget vector f_t, then the output from the input gate gets point-wise added. This step updates the cell state to new values that the neural network finds relevant:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C_t}. \tag{4.11}$$

In the end, the output gate decides what the next hidden state should be. The previous hidden state and the current input are passed into a sigmoid function. The newly modified cell state is passed to the tanh function and then multiplied by the output of the sigmoid gate o_t. The output is the new hidden state h_t. The new cell state and the new hidden state are carried over to the next time step.

$$o_t = \sigma(W_o[h_{t-1}, x_t]) + b_o. \tag{4.12}$$

$$h_t = o_t * tanh(C_t). \tag{4.13}$$

In all these formulas, W are the weight matrices and b are the bias vectors.

4.5 K-Fold cross validation

The K-fold cross validation is a technique to test the model ability to predict new data, aimed at preventing problems such as overfitting. The final goal is to reduce the dependence of the model on the data, which manifests itself by low error on the training set but a large error on data excluded from the training.

The letter K indicates the number of subgroups in which data are split. Typically, $K = 5$ or $K = 10$ as these values have been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance. Figure 4.17 shows a graphical example of data splitting.

Data in the training set are labeled as the belonging class and are used to train the model; instead, data in the test set are considered as new data and are used to test the trained model.

The algorithm of k-fold cross validation in pseudo-code is:

1. divide the dataset into K groups.

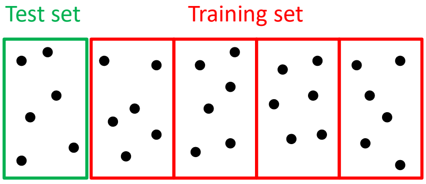


Figure 4.17: In the 5-fold cross validation the dataset is divided in five groups, four for the training set and one for the test set.

2. For each group:
   (a) train the model using $K-1$ groups as training set.
   (b) Test the model using the remaining group.
   (c) Compute the accuracy and store it in an accuracy list.
   (d) Repeat the process until every K-fold served as the test set.
3. The final accuracy is the average accuracy, and is obtained as

$$Accuracy = \frac{\sum\limits_{Acc \in L_A} Acc}{K},$$

where L_A is the accuracy list. Figure 4.18 depicts the scheme of K-fold Cross Validation.

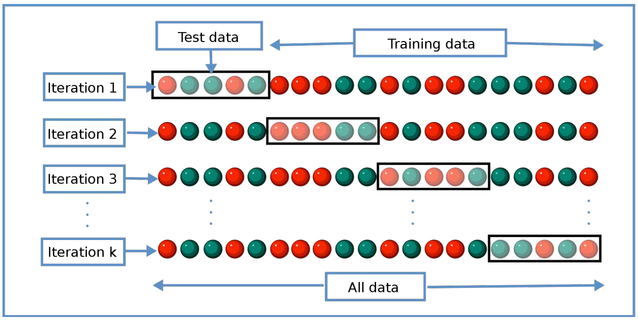


Figure 4.18: Scheme of K-fold Cross Validation.

Chapter 5

Application of diagnostic imaging techniques

This chapter discloses the studies in the context of image processing. First, a goal-driven image segmentation method is introduced, with an application regarding the segmentation of carpal bone in MR images of patients with RA. Then, the issues related to raters' agreement is addressed. An annotation tool for supporting the annotation process is introduced, along with a very first study on experts' concordance. Finally, a pipeline for an unsupervised evaluation of segmentation results is presented.

5.1 Unsupervised segmentation of wrist bones

The goal-driven approach to segmentation

The present Section is devoted to the definition of the goal-driven formulation of image segmentation here proposed. Let $I \subset \mathbb{Z}^2$ the pixel lattice, $\mathcal{X} = \{x_i\}_{i \in I}$ the collection of the observed pixel intensities, and let the image be formally represented as $X = I \times \mathcal{X}$. With this formalization, each individual pixel is conveniently associated with the pair (i, x_i) so that both the spatial location $i \in I$ within the imaged area and the intensity x_i are specified.

According to the classical definition [6], the region segmentation of an image X into disjoint non-empty subregions $X_1, \ldots, X_M$ has to satisfy the following requirements:

i) $\displaystyle\bigcup_{j=1}^{M} X_j = X$;

ii) X_j is connected for each $j = 1, \ldots, M$;

iii) $P(X_j) = \texttt{true}$ for each $j = 1, \ldots, M$;

iv) $P(X_j \cup X_k) = false$ for each $j \neq k$, where P is a predicate of homogeneity. In order to focus this classical definition with particular emphasis on the goal to be achieved, one can consider a goal-driven definition of image segmentation, which is aimed at extracting some ROIs which are interesting for the application itself. Here, instead of searching for an image partition, the goal is to split the image X into a finite collection of disjoint ROIs $X_1, \ldots, X_M$ and a background (also denoted as reject-region) $B = \left(\bigcup_{j=1}^{M} X_j\right)^c$, where $(\cdot)^c$ denotes the complementary set. Therefore, each ROI has to satisfy the following requirements:

i) $\bigcup_{j=1}^{M} X_j = X \setminus B$;

ii) X_j is connected for each $j = 1, \ldots, M$;

iii) $P(X_j) = \texttt{true}$ for each $j = 1, \ldots, M$;

iv) $P(X_j \cup X_k) = false$ for each $j \neq k$;

v) $P(B) = false$,

where P is a goal-driven predicate of homogeneity, while the background B does not satisfy the homogeneity predicate (when B is not connected, all the partitions do not satisfy the predicate). Here, the goal-driven predicate P encompasses the information related to the topological, morphological, and statistical properties of data, and depends on the user-defined goal to be achieved.

The predicate is defined by the user and reflects their expertise (e.g., a medical doctor, a photo-interpreter, an ultrasound technician, etc.). Indeed, as in the actual practice and with particular emphasis on the medical domain, it is important to define what is the aim of the proposed procedure. The presented method aims at emulating such an approach by including some prior-knowledge expressed in terms of goal to be achieved. Such a goal is formally expressed in mathematical terms, and is deployed by coding the corresponding control functions.

Within the goal-driven definition of image segmentation, the function P literally drives the following processing. In particular, the goal is defined as a collection of properties that are desired for the ROIs and not for the background. These properties are determined by the application for which the segmentation task is addressed. Among the possible options, in this paper, two properties are always deployed. They refer to the intensity and the distribution of the pixels within each ROI. Indeed, the former is related to some visual features of the regions to detect (e.g., bright structures in MR), while the latter is aimed at properly modeling the distribution of the data within the regions taking into account the type of data under analysis (e.g., modeling the type of noise affecting the available acquisition). In addition, other properties such as some positional, topological, or geometric ones, can be integrated in the proposed framework.In the formulation of the hypotheses, the image intensity X is modeled as a realization of a two-dimensional stochastic process, i.e., the intensity x_i of the i-th pixel ($i \in I$) is modeled as a random variable.

To formally define the aforementioned goal, let h_g and h_s be the two hypotheses, i.e., the pixel-intensity property and the statistical property, respectively. The former is defined with respect to a set $A \subseteq X$ as:

$$h_g(A) = \texttt{true} \iff \begin{cases} \mathbb{E}\{x_i\} \in J_m, \\ \mathrm{Var}\{x_i\} \in J_v \end{cases} \forall (i, x_i) \in A, \qquad (5.1)$$

where J_m and J_v are two intervals that specify which values of the sample-mean and the sample-variance are deemed feasible.

Then the latter hypothesis addresses the statistical properties of the ROIs. Indeed, the set $A \subseteq X$ is said to satisfy hypothesis h_s if the samples of pixel intensities x_i, when conditioned to their membership to the same region, are independent and identically distributed (i.i.d.) and their probability density function (pdf) belongs to a given parametric pdf family f (e.g., Gaussian, Gamma), i.e.:

$$h_s(A) = \texttt{true} \quad \iff \quad x_i \sim f \;\; \text{i.i.d.} \quad \forall (i, x_i) \in A, \qquad (5.2)$$

In addition, other properties that may be considered can be similarly defined. For instance, a positional property can be considered if the application requires that the target regions are located roughly in a certain area of the image (e.g., in the center of the image). Such property, defined h_p, is considered in the application to MR images, which requires seeds to be in the center of the image, thus avoiding searching for bones close to the image borders, and is defined as:

$$h_p(A) = \texttt{true} \quad \Longleftrightarrow \quad (i, x_i) \in \overline{X}_F, \quad \forall (i, x_i) \in A, \tag{5.3}$$

with $\overline{X}_F$ being the central area of the image X (i.e., without the border of size F).

Now, let H be the collection of properties to be satisfied:

$$H = \{h_g, h_s, h_{other}\}, \tag{5.4}$$

with h_{other} collectively indicating possible optional properties that can be specified by the user via appropriate mathematical expressions. Then, let P be a function from the power set of the image X to the Boolean set:

$$P : 2^X \rightarrow Bool \text{ so that } P(A) = \bigwedge_{h \in H} h(A) \tag{5.5}$$

The predicate and its components are used to guide the behavior of the presented segmentation method, but it does not ensure that the outputs satisfy all the conditions. Nevertheless, P drives the computation, i.e., it is used to encode the goal within the processing steps and it is applied for specific operational needs. In particular, due to the seed-growing nature of the proposed approach, the predicate will be conveniently applied to either the seed points or the ROI samples.

Overview of the proposed approach

The proposed segmentation method is aimed at extracting, in an unsupervised fashion, elements of interest from 2-D images, taking into consideration the user-defined goal and the application domain to which it is applied.

The resulting output would be a label image $L = \{l_i\}, i \in I$, defined over the same (pixel) lattice, whose pixels are associated with labels identifying a particular element, i.e. $l_i \in \Lambda$, with $\Lambda = \{0, 1, 2, \ldots, M\}$ and M being the number of different elements, i.e., of the different ROIs, and the label $l_i = 0$ standing for a background pixel.

The rationale of the method is to combine weighted graphs, parametric density modeling, and Markov random fields to benefit from both the topological and the statistical properties of the input image in the generation of an output map made of the regions of interest to the application. The regions to be considered in the final partition are obtained from an initial finite set of seeds S_0, placed in the image according to a random grid. Nevertheless, based on the definition of the goal, a subset of S_0 is selected, and such remaining seeds are used to identify the candidate ROIs.

Here, the proposed method acts driven by the predicate P defined in the previous Section, even though it does not need to be verified on the final segmentation obtained. With this in mind, the hypotheses h_g and h_p are applied with regards to the seeds.

In particular, first of all h_g is applied to $\tilde{X}_s$, with s being a candidate seed point, $\tilde{X}_s$ a small region around s identified via a windowing operator. The window prevents the selection of an isolated noisy pixel as one of the seed points. For the experiments of this paper, it is set equal to 5×5.

The goal-driven formulation of the proposed method is indeed guided by the predicate, which is used as a filtering operator on the initial set of seeds in the random grid. It drives the subsequent collection of ROI samples, as disclosed in the following.

Then, the predicate h_s is applied to the set of samples $X_n \subset X$ that are extracted via the graph-based computation (the formal definition will follow). Differently from the selection of the seeds, the choice of the pdf does not directly imply the elimination of some candidate ROIs in the proposed method. Nevertheless, the samples that poorly fit the chosen model will provide a negligible contribution to the energy function of the Markovian formulation of the problem, thus favoring that the candidate ROI is removed from the final map.

Finally, in case the predicate h_p is required, e.g., in the application to MR images, the argument of such a function is $X_n \subset X$, thus requiring the ROIs to be placed in the specific portion of the image.

The preliminary stage of the proposed approach ends once the goal, and so the P function, are defined. Then, the three phases of the proposed method start, whose flowchart is shown in Figure 5.1.

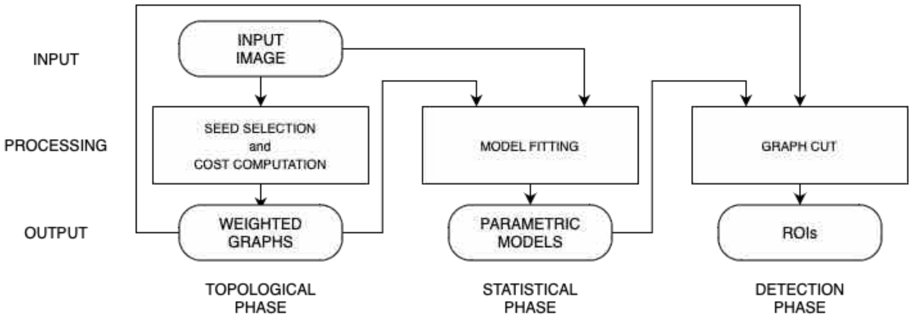


Figure 5.1: Method flow chart

Following the goal-driven philosophy of the method, the seeds in the initial set S_0 are filtered according to the aforementioned predicate. Specifically, the properties of H referring to the seed-selection phase are verified for each seed $s \in S_0$, and s is retained if and only if they are `true`. In the following, the subset $S \subset S_0$ of the original set of seeds that are selected according to P is denoted as $S = \{s_n\}, n = 1, 2, \ldots, N$.

Then, the propagation algorithm in [175] is used to compute a set of cost functions $F = \{F_n\}$, with n identifying each one of the seeds $s_n \in S$. Each cost function is defined over the whole image such that $F_n : I \to \mathbb{R}$ (Phase 1). Differently from the graph-based method mentioned in the introduction [24, 26, 27, 28, 29], no graph-cut is actually performed at this level. Hence, no segmentation is obtained in the first phase.

On the contrary, a set of samples $X_n \subset X$, defining a region associated with the particular seed s_n, is extracted from the image based on each cost function $F_n, n = 1, 2, \ldots, N$. They can be considered as fuzzy syntactic primitives whose properties depend on the application domain, and that influence the homogeneity predicate. The complementary region (i.e. the set of image pixels not assigned to any set of samples X_n) is defined as the initial background $B_0 = X \backslash \{\cup_{n=1}^{N} X_n\}$. Moreover, according to the graph-based processing, it is not forbidden for the same pixel to be assigned to multiple sets (e.g., when it occurs that two random seeds are close and inside the same region). Nevertheless, the Markovian framework characterizing the further processing steps is aimed at taking care of such a situation.

Finally, a set of parametric models, each one associated with each seed $s_n \in S$, is estimated using the samples collected in the regions X_n (Phase 2), as described in detail in Section 5.1.

The segmentation problem is then formulated in a Bayesian probabilistic graphical framework by defining a Markov random field model. Accordingly, MAP decision rule is formulated as the minimization of a suitable energy function [176]. The energy function is composed of two terms, a unary term $D_i(\cdot)$ and a pairwise term $V(\cdot)$ (see Equation 5.6). The unary pixel-wise contribution is related to: i) the parametric model associated with the set of samples corresponding to each seed; ii) a background-specific potential associated with the set of samples in the initial background B_0; and iii) a spatial feature related to the distance, in the image lattice, between the location of a given pixel and the location of each seed. Conversely, the pairwise contribution brings about local-contextual information. Indeed, the energy function is defined as:

$$U(L|X) = \sum_{i \in I} D_i(x_i|l_i) + \beta \sum_{i \sim j} V(l_i, l_j), \tag{5.6}$$

where $i \sim j$ denotes pixels i and j being neighbors according to a first-order neighborhood system, $i, j \in I$ (Phase 3).

Because of the Bayesian formulation of the problem, it is possible that the final label map L does not contain all the labels from the seeds list. It is worth recalling that each seed is associated with a particular label from the preliminary set of labels $\Lambda' = \{0, 1, 2, ..., N\}$. Indeed, this filtering behavior is actually sought for, as it allows the proposed method to autonomously remove, in the output result, the contributions of spurious seeds (i.e. those seeds that do not correspond to well-defined regions). Additionally, such a Bayesian formulation also solves possible conflicts, as for example the cases where multiple seeds are placed inside the same region. The minimization process assigns a unique label to such a region. In other words, it allows the method to automatically determine the number of most relevant regions, whereas N plays the role of an upper bound on such number.

To conclude, according to the novel definition of goal-driven segmentation, the resulting final map L is composed by the extracted ROIs X_j $(j = 1, 2, \ldots, M)$ and the background B.

The method

Graph-based Cost Computation The graph-based approach proposed here is an extension of the method described in [175] and is based on the computation of the set of cost functions $F = \{F_n\}$, each one associated with a specific seed $s_n \in S$.

For the sake of simplicity and ease of notation, this section describes the cost computation process focusing on a specific seed, provided that the extension to the case of multiple seeds is straightforward and only requires to replicate the process for each seed. We will therefore report just the index n without specifying $n = 1, 2, \ldots, N$. In addition, the input image is assumed to be scalar valued, while the extension to vector valued images is straightforward and only requires some minor mathematical adjustment.

A grid is defined over the pixel lattice and each vertex of the grid is randomly shifted in the vertical and horizontal directions, according to a uniform distribution. The resulting vertices correspond to the chosen seeds. The spacing of the grid can be considered as hyperparameter depending on the size of the object of interest. On the one hand, a small spacing ensures that a seed is placed in each object, but yields to many computations. On the other hands, large spacing may lead to neglecting some objects of regions. Since the seed selection process deletes redundant seeds, small spacing is in general preferable than large spacing. In the experimental phase, based on the size of the ROIs, such a hyperparameter was assumed to be the 10% of the image width. Small changes of the spacing did not change the final output from the visual point of view.

Let the image X be mapped into a non-empty, fully-connected, undirected and vertex-weighted graph $G = (I, E)$ where I is the set of vertices[1] representing the pixels in the image and $E \subseteq \{(i, j) \in I \times I \mid i \neq j\}$ is the set of edges. The goal of this section is thus to describe the computation of $F_n(i), \forall i \in I$, given the seed $s_n \in S$.

---

[1] for ease of notation, here vertices are indicated as letters without ordering subscript, differently from Section 3.1.

For each couple of nodes i and j, let w be a difference function such that:

$$w : I \times I \rightarrow \mathbb{R},\ w(i,j) = |x_i - x_j|, \tag{5.7}$$

where x_i is the intensity of pixel $i \in I$, and A and B are the minimum and maximum admissible differences. An example is given by $[0, 255]$ in the case of 8-bit gray level images. The function in Eq.5.7 represents the difference between each pair of vertices i and j. The higher the value of $w(i,j)$, the more different the two nodes are in terms of gray level.

By fixing a specific node s_n in the graph (i.e., the seed), each point in the image, and thus each node in the graph, can be associated with a value representing its dissimilarity with respect to s_n. Indeed, it is possible to compute the difference function with respect to the seed according to $w_n(i) = w(i, s_n)\ \forall i \in I$. The proposed cost computation process is based on the computation of $w_n(i)$ for each node in the graph and according to a flooding scheme [24].

At the beginning, all of the nodes are in the *unvisited* status except for the seed, whose cost is zero. A set T, initially containing only the seed node, is defined to keep track of all the visited nodes. Then, the flooding process is started from the seed and toward the connected neighbors (a 4-connected neighborhood model is chosen). Among the neighboring nodes, the one providing the minimum value of w_n is chosen and added to T. Consequently, the value of the cost function associated with the considered node is computed. Then the flooding process is iterated until the set T of visited nodes coincides with the whole set of nodes I (see Figure5.2.)

Concerning the computation of the cost function, at each iteration two nodes are taken into consideration: i) the node i being considered; and ii) its father node i^-, which is the node that, according to the flooding scheme, led to i. In particular, if the weight $w_n(i)$ of i with respect to the seed is larger than the cost of i^-, then the cost associated with i is set equal to $w_n(i)$, otherwise the cost value of $w_n(i^-)$ is inherited (see Equation 5.8). In this way, the cost function is never decreasing with the flooding scheme:

$$F_n(i) = \begin{cases} w_n(i) & \text{if } w_n(i) > F_n(i^-), \\ F_n(i^-) & \text{if } w_n(i) \leq F_n(i^-). \end{cases} \tag{5.8}$$

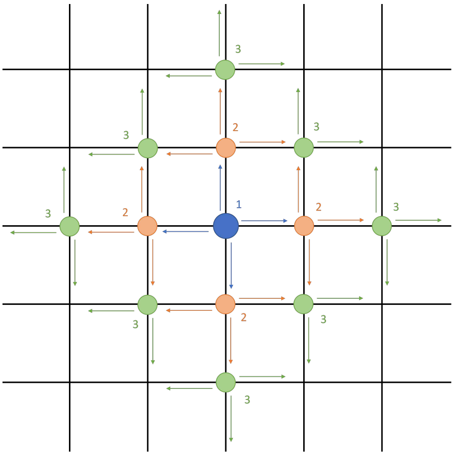


Figure 5.2: Schematic representation of the flooding process starting from the seed node (blue circle) and using a 4-connected neighborhood model.

Leveraging on such propagation algorithm, a cost value is assigned to all nodes and the process defines an optimal path from the seed to each vertex. To better understand the process, Figure5.3 shows the assignment of such cost values to all the nodes of the graph built on top of a magnetic resonance image. Looking at the images from left to right and from top to bottom it is possible to appreciate the assignment of cost values starting from low values and moving to high values. The nodes assigned with a cost value are painted black, and the figures are subsampled (i.e., for the sake of brevity, not all the steps are reported). It is interesting to visually verify that, moving away from the seed, the values are always increasing. Such a procedure yields to the set of region samples X_n associated with the n-th seed. In particular, the cost function is thresholded according to the method in [175]. The set of pixels associated with costs lower than such a threshold are assigned to the set X_n. The resulting region is not disjoint due to the non-decreasing formulation of the cost function F_n.

As mentioned at the beginning of this Section, repeating the process for each seed s_n yields to the set of region samples $\{X_n\}, n \in \{1, 2, \ldots, N\}$. Indeed, due to the definition of the graph and the cost function in Equation 5.8, such regions are characterized by

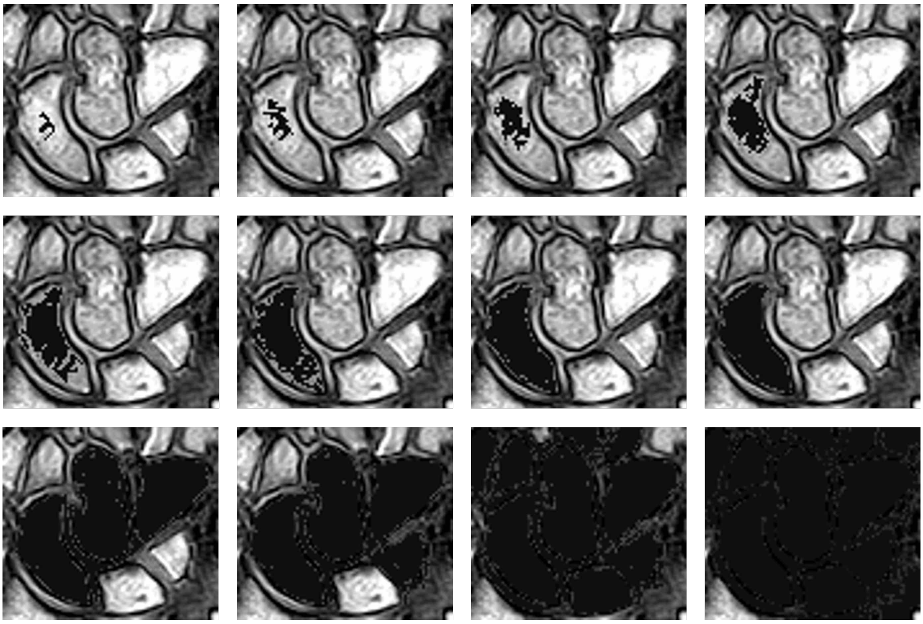


Figure 5.3: Assignment of the cost values to the nodes in the graph. The figures from left to right and from top to bottom show the assignment of such costs from lower to higher values. The nodes assigned with a cost value are painted black, and the figures are subsampled (i.e., not all the steps are reported). The reader may notice that, moving away from the seed, the values are always increasing.

homogeneity and low granularity. The set of regions $\{X_n\}, n \in \{1, 2, \dots, N\}$ obtained from the present phase of the method generally does not cover the whole image, i.e., it leaves out the initial background defined as B_0.

Parametric Model Estimation

The cost computation phase is followed by the parametric model estimation, which can be deployed by means of the methods presented in Section 3.3. The observations in each set of samples X_n, $n \in \{1, 2, \dots, N\}$ are considered, focusing on the radiometric information contained in $\mathcal{X}$, and not on the whole image X conveniently defined as observations and pixel locations.

Energy Minimization through Graph Cut In the proposed method, the energy function in Equation 5.6 is composed of a unary term and a pairwise term. The unary pixelwise term $D_i(x_i|l_i)$ is defined differently in case it refers to the seed-specific labels (i.e., $l_i = 1, \dots, N$) or to the background B_0 (i.e., $l_i = 0$). In the former case it is related to the seed-specific parametric models and to a term measuring the distance between the considered pixel x_i and the seeds. Conversely, in the latter case, it is related to a piecewise constant background-specific potential. The pairwise potential is defined as the Potts model [176]. Indeed, the two terms are as follows:

$$\begin{aligned} D_i(x|n) &= \begin{cases} d_i(x|n) & n = 1, 2, \dots, N \\ c_1 & n = 0, i \in B_0 \\ c_2 & n = 0, i \in I - B_0 \end{cases} \\ V(n, m) &= 1 - \delta(n, m), \end{aligned} \tag{5.9}$$

where $\delta(v, w)$ represents the Kronecker delta, whose value is 1 if and only if v equals w, and zero otherwise, $d_i(x|n)$ is the seed-specific potential, and c_1 and c_2 are the values of the piecewise constant background-specific potential.

Concerning the seed-specific potential, let $\hat{p}(\cdot|n)$ be the PDF estimate obtained as described in the previous section from the set of samples $X_n, n \in \{1, 2, \dots, N\}$ and according to either a Gaussian or a Gamma model. Then, the potential can be written as:

$$d_i(x|n) = -\ln \hat{p}(x|n) - \gamma \left[\Delta\left(i, s_n\right)\right]^{-1}, \tag{5.10}$$

where $\Delta(a, b)$ is the Euclidean distance between pixel locations a and b in the image plane ($a, b \in I$), and γ is a positive coefficient balancing the two contributions.

Concerning the piecewise-constant background potential, the two values c_1 and c_2 are automatically chosen according to the p-th and $(1-p)$-th percentiles of the distribution of the seed-specific potentials $d_i(x|n)$. The rationale is to balance the background-specific and the seed-specific potentials so as to favor the goal-driven result while also not censoring any possible outcome in the output label map Y. Both the parameters γ (Equation 5.10) and β (Equation 5.6) are determined via a trial-and-error procedure.

In particular, the term $\Delta(i, s_{l_i})$ represents the spatial distance between the location of the pixel $i \in I$ in the image lattice and the location of the seed s_{l_i} that corresponds to its label l_i. While the first component of the unary potential is traditionally related to the likelihood of the data, the integration of the second contribution in the unary term is inherited from a family of segmentation methods that explicitly integrate a spatial feature in the formulation of the unary term [177]. This is also responsible for linking the intermediate statistical procedure with the initial topological cost computation, and the final cut yielding the results.

The minimization of a Markovian energy function, like the one defined in Equation 5.6, has been dealt with using several techniques in the last few decades. Here the graph cut approach [142] is used, which is based on the reformulation of the energy minimization problem as a max-flow/min-cut problem over a suitable graph. In the case of binary labeling, graph cut approaches are also proven to converge, in polynomial time, to the global minimum. In the case of more than two labels, ad-hoc techniques have been formulated. Here, the chosen technique is the alpha-beta swap technique, which reformulates the problem as a sequence of binary sub-problems and, for each sub-problem, a global energy minimum is reached through the max-flow/min-cut formulation. Convergence to a local minimum with strong optimality properties is guaranteed in this case [142].

Experimental results

The proposed method was tested on ten MR T1-weighted images from the database in [7]. They are images of the wrist district depicting seven carpal bones (i.e. capitate, hamate, lunate, scaphoid, trapezium, trapezoid, and triquetrum), in the coronal plane. The acquisition was performed via the 0.2 Tesla Artoscan (Esaote Spa, Genova, Italy). The dimension of each image is 80 × 100 pixels, the maximum gray-level is 255, and the ground truth segmentation was performed by medical operators with extensive experience in RA evaluation. The goal to be achieved is the detection of bright regions corresponding to the trabecular part of the wrist bones.

Specifically:

$$h_g(A) = \texttt{true} \iff \begin{cases} \mathbb{E}\{x_i\} \in (m_{MR}, M_{MR}), \\ \mathrm{Var}\{x_i\} \in (0, v_{MR}) \end{cases} \forall (i, x_i) \in \tilde{X}_{s_n}, \tag{5.11}$$

with m_{MR}, M_{MR}, and v_{MR} being thresholds that can be set either directly by the user (e.g., the expert selecting those values based on prior knowledge) or automatically, e.g., in a supervised fashion. Here, the former strategy has been adopted;

$$h_s(X_n) = \texttt{true} \quad \iff \quad x_i \sim \mathcal{G} \ \text{ i.i.d.} \quad \forall (i, x_i) \in X_n, \tag{5.12}$$

with $\mathcal{G}$ being the Gaussian distribution;

$$h_p(X_n) = \texttt{true} \quad \iff \quad (i, x_i) \in \overline{X}_F, \quad \forall (i, x_i) \in X_n, \tag{5.13}$$

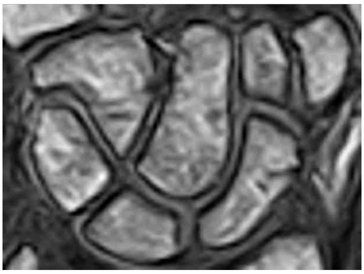
where F was chosen as the 10% of the smallest image size.

Figure 5.4 shows one of the original images, the corresponding ground truth, and the segmentation maps obtained by the proposed method and by other four state-of-the-art techniques.

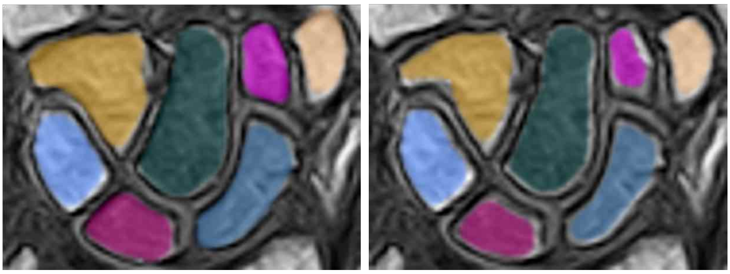
Based on the ground truth in [7], the results are compared with those obtained by the active-contours method in [179], the graph-based method in [175], the iterative spanning forest (ISF) [29], and with a classical MRF model (i.e., without the term $[\Delta(i, s_{l_i})]^{-1}$ in Equation 5.9), taking into account the *dice score* as performance indicator [178].

As aforementioned in Section 5.1, the method in [175] is the baseline for the current work. Both the graph-based approach and the propagation mechanism are in common, but no Bayesian formulation is included. As a consequence, each ROI is extracted one at time, or an optimal criterion for the seeds placement has to be defined. In particular, the segmented bones are smaller then the ones obtained via the proposed method: the false positive pixels are reduced through the developed approach.

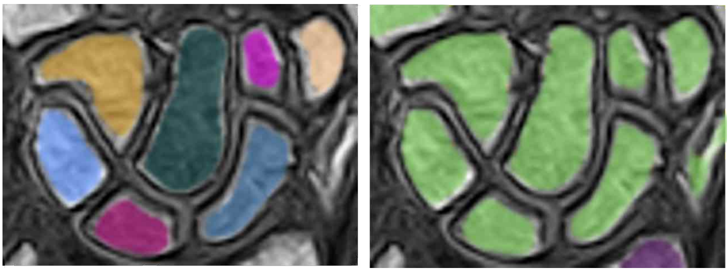
The other starting point for the proposed method is an MRF model in its classical formulation, i.e., without the distance term which is present in Equation 5.9. Here, not only the carpal bones



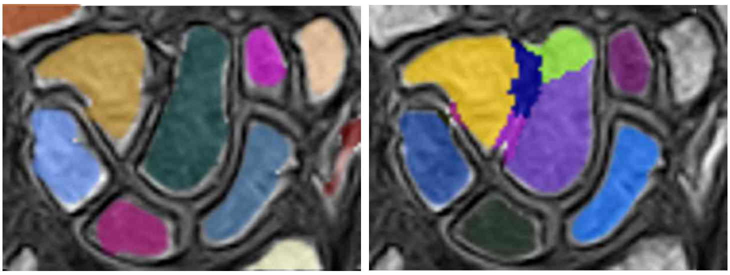
(a) Original



(b) Ground truth (c) Proposed method



(d) Graph-based method [175] (e) Classic MRF [176]



(f) Active contours [179] (g) ISF [29]

Figure 5.4: Carpal bones in MR: 5.9a original data; 5.4b ground truth [7]; 5.4c results obtained by the proposed method; 5.4d results obtained by the baseline graph-based method in [175] (comparison 1); 5.4e results obtained by the classical formulation of MRF models (comparison 2); 5.4f results obtained by the method in [179] (comparison 3); 5.4g results obtained by the method in [29] (comparison 4).

are extracted, but also other regions. Furthermore, one can notice how the distance term in Equation 5.9 allows for labeling each bone with a different label, meaning that each ROI is represented by a different seed.

Then, the method in [179] is an example of a deformable model. It consists of a flexible active contours model which makes use of thresholding or gradient-based methods. Even though it is widely used in the medical field, in this case the shapes of the segmented ROIs result to be less regular and compact than the ones obtained by the proposed method. Moreover, the method in [179] also provides a significant amount of false positive pixels, resulting in a set of small and spurious ROIs.

Finally, the method presented in [29] is designed for superpixel computation, and hence it results in an over-segmentation of the image. Indeed, in many cases, the bones were over-segmented. Thus, in order to enable the comparison with the proposed method, the output superpixels have been manually merged. In some cases, the method in [29] was not able to detect some of the carpal bones, thus affecting the final mean scores. As for the quantitative comparison, the accuracy measured by using such method does not mirror its actual capability. Indeed, the performances are positively biased by the manual merging operations that have been performed on the output segmentation to allow the comparison.

The performance measures reported in Table 5.1, together with the qualitative results reported in Figure5.4, denotes the capability of the proposed method to provide a segmentation result characterized by regular shape, achieving effective performances which are also consistent across the different bones, and avoiding spurious and noisy ROIs in the output map. The regularization of the MR artifacts and noise is guaranteed by the MRF model, that allows segmenting all of the bones, which are present in the considered image, as connected components that can be easily extracted and identified. The values in bold were statistically tested to significantly outperform their rivals the (t-test, $p < 0.01$).

Table 5.1: DICE score comparison.

	Proposed Method	**Graph-based method [175]**	**Classic MRF [176]**
Capitate	**0.903 ± 0.032**	0.826	0.877 ± 0.036
Hamate	**0.875 ± 0.037**	0.785	0.843 ± 0.078
Lunate	**0.850 ± 0.050**	0.765	0.819 ± 0.039
Scaphoid	**0.877 ± 0.027**	0.789	0.842 ± 0.034
Trapezium	**0.880 ± 0.037**	0.832	0.735 ± 0.174
Trapezoid	0.799 ± 0.086	**0.813**	0.740 ± 0.067
Triquetrum	**0.911 ± 0.031**	0.695	0.864 ± 0.031
Total	**0.871**	0.786	0.817
	Proposed Method	**Active contours [179]**	**ISF [29]**
Capitate	**0.903 ± 0.032**	0.878 ± 0.045	0.830 ± 0.078
Hamate	**0.875 ± 0.037**	0.874 ± 0.046	0.850 ± 0.049
Lunate	**0.850 ± 0.050**	0.826 ± 0.057	0.737 ± 0.117
Scaphoid	**0.877 ± 0.027**	0.857 ± 0.034	0.805 ± 0.084
Trapezium	**0.880 ± 0.037**	0.854 ± 0.068	0.482 ± 0.289
Trapezoid	0.799 ± 0.086	0.792 ± 0.062	0.614 ± 0.295
Triquetrum	**0.911 ± 0.031**	0.900 ± 0.037	0.665 ± 0.296
Total	**0.871**	0.854	0.712

5.2 Annotation tool

The development of a tool facilitating the annotation allows to widen the pool of annotators, enabling the creation of datasets with more different labels and therefore more robust and reliable. Also, a web service providing such a tool would enhance the accessibility by raters, thus granting a large number of annotations.

For such reasons, aimed at defining a numerical-segmentation-driven procedure for annotating medical figures, a pre-processing phase based on superpixels segmentation is proposed. The medical experts are required to select the superpixels of interest in an interactive fashion, and then to manually correct the annotation.

Figure 5.5 depicts the phases of such operation.

The original data is over-segmented via a superpixel algorithm, producing a grid to be superimposed on the data. An underlining labels image is produced, to enable superpixels annotations. The final output consists of the binary annotation by the expert.

Such a procedure is supposed to reduce intra- and inter-rater variability and the dependence on the skill of the annotators, thus enabling more robust classifiers, trained on such labeled data. In addition, a preliminary evaluation on the possibility that the presented pipeline can reduce time needed for manual annotation was

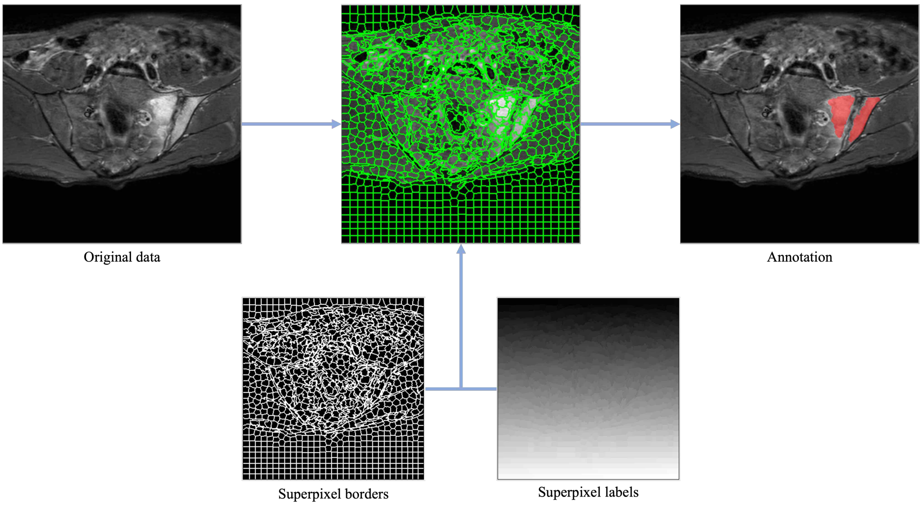


Figure 5.5: Workflow of superpixels-driven annotation.

conducted. Five raters with different levels of experience annotated some images both with and without the use of the annotation tool. The time required for the annotation reduced with the support of the pre-processing. While this is still a conjecture rather than a result, future study will be able to confirm such statement.

5.3 Raters agreement

A preliminary phase in evaluating raters agreement in annotations is here disclosed. It is aimed at studying the possibility to reduce inter-rater variability and the dependence on the skill of the annotators in the labeling phase. Hence, five annotators were asked to mark every region suspicious of the presence of bone marrow edema regardless of its nature (inflammatory, degenerative, traumatic), into five STIR images of the sacroiliac joints, as in [76, 180]. Bone marrow edema is a hyperintense areas within the bone marrow on STIR sequence. In general, lesions manifest themselves as regions of STIR images which are brighter then the signal within the reference region, i.e., the region where it is unlikely that inflammations occur. Even though this is a precise definition, it is rarely rigorously applied in practice. In fact, medical doctors are likely to consider lesion a region which is brighter than the neighboring ones. As a result, the manual segmentation of lesions is very subjective.

In the present experiment, raters annotated the images via two procedures, i.e., manually, and interactively by using the annotation tool described in Section 5.2. The latter delivered the pre-segmentation results based on the superpixel methods described in Section 3.4, i.e., LSC and TPS. The agreement among raters is evaluated in terms of DICE coefficient [178] and Intersection over Union (IoU) coefficient [181]. Such indices are computed by coupling raters and the average values are reported in Table 5.2. Also, the global agreement is computed as average indices over the five considered cases.

Table 5.2: Raters agreement.

	DICE coefficient			IoU coefficient		
	Manual	LSC	TPS	Manual	LSC	TPS
Case 1	0.89	0.98	0.99	0.80	0.95	0.98
Case 2	0.86	0.97	0.96	0.75	0.95	0.92
Case 3	0.74	0.98	0.98	0.60	0.96	0.95
Case 4	0.88	0.97	0.97	0.80	0.94	0.94
Case 5	0.78	0.96	0.98	0.64	0.92	0.95
Average	0.83	0.97	0.97	0.72	0.94	0.95

In addition, Figure 5.6 depicts an example of raters agreement visualization, based on the label map. It represents a concordance heat-map where warm colors represent higher agreement among the five raters (i.e., in descending order red, yellow, green, light blue, and blue). It appears that the major issues arise close to the borders of regions, and are highlighted as they are less defined. It is noteworthy that in all the cases, the raters agreement increased when considering the superpixel-based pre-annotation. Indeed, the preprocessing phase enable an enhanced visualization of borders thus yielding lower variability of annotation, acting as a standardizing procedure.

By leveraging on robust annotations, supervised classifiers are supposed to benefit in terms of performance and both classification and predictive ability. The proposed procedure could thus address

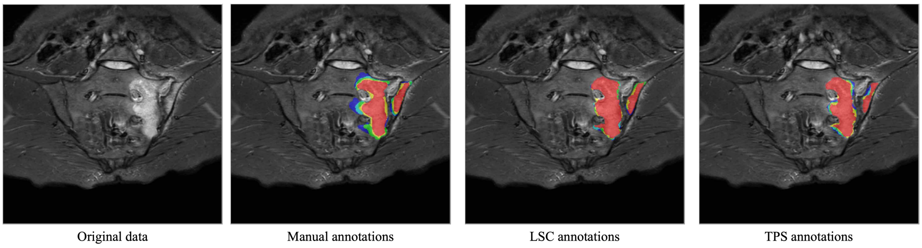

Original data Manual annotations LSC annotations TPS annotations

Figure 5.6: Example of raters agreement visualization.

a significant problem of supervised method, tackling the problem of poor ground truth data. Of course, any unsupervised method could be deployed in the pre-processing phase; the choice depends on the application domain.

The proposed pipeline is, in fact, a solution bridging unsupervised and supervised approaches, aimed at better fulfilling the final semantic task, and can be applied not only in the biomedical context, but also in any framework where raters agreement may affect the quality of the results.

5.4 Unsupervised segmentation evaluation

In the present section, a method for unsupervised evaluation of image segmentation results is presented. The description follows the dissertation in [182].

The main objective of the proposed work is to evaluate the goodness of a segmentation result in an unsupervised fashion, namely, based on local contrast information related to the interface between the internal and external voxels of a ROI. In particular, the aim is to evaluate the goodness of a segmentation result with regards to a specific application, which here is the adherence of the segmentation to the cortical-trabecular interface of bones, thus fostering subsequent analysis such as the detection of relevant landmarks. It is then clear how a reliable gradient measure is the basis of the whole processing chain. Intuitively, the proposed gradient measure evaluates the variation of the gray levels across the interfaces, thus defining an approximation of the gray levels gradient on the border of the segmented object. Since the best location of the interface between the inner and outer voxels of a ROI is associated with the local maximum contrast, the comparison between two different segmentations is based on the statistical analysis of the local gradient values. Given a bone segmentation result, the boundary voxels are extracted to analyze gray levels information in a small local neighborhood. Instead of directly applying the original MRI gray level values, some pre-processing is proposed. Morphological dilation and erosion operations are independently performed on the raw volume, in order to spatially expand bright and dark areas, respectively. These two volumes are used separately as input to the gradient calculation phase. The difference between the gray levels in the two processed volumes is mapped onto the voxels on the border, representing the volume boundary, hence the interface between cortical and trabecular bone. The resulting volume boundary texture is the difference after dilation and erosion; the higher the value of the resulting texture, the more different are the tissues inside and outside the volume. For the comparison of segmentation results, a fuzzy criterion based on local statistics is defined.

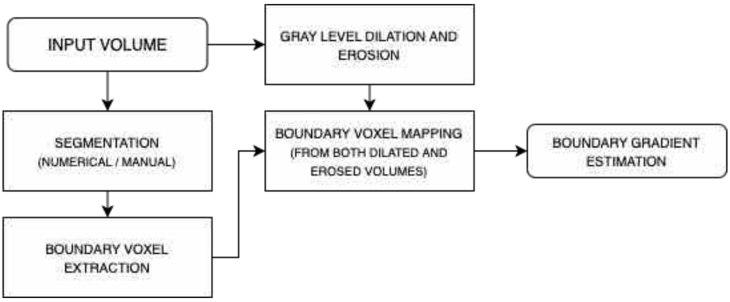


Figure 5.7: Flow chart of the proposed approach.

The method

Segmentation The starting point of the proposed method is volume segmentation. In particular, here both manual and numerical segmentation are considered. As for the manual segmentation, it was performed by medical experts in MR volumes evaluation, by using the tool RheumaSCORE (Softeco Sismat S.r.l.). On the other hand, the automatic approach refers to the method in [24], which is a graph-based, unsupervised, and adaptive method, which therefore does not require any a-priori knowledge.

Boundary voxel extraction From the segmented binary volumes, the voxels on the border are extracted in a classical fashion, by leveraging on binary morphological operations. In particular, two sets of boundary voxels are defined. More specifically, let V be the binary segmented volume and let s be a cubic structuring element of dimension $3 \times 3 \times 3$, and whose entries are equal to 1. The sets of boundary voxels are obtained via the equations:

$$\textsc{border}_1 = V - (V \ominus s); \; \textsc{border}_2 = (V \oplus s) - V, \tag{5.14}$$

where $\ominus$ and $\oplus$ are the binary erosion and dilation operators respectively [160].

Note that $\textsc{border}_1$ puts more emphasis on the gray levels just inside the volume, while $\textsc{border}_2$ refers to the gray levels just outside the volume. It is worth mentioning that no surface mesh has been considered. Indeed, the nodes of the mesh are an approximation of the volume surface, which yields to the loss of some information content. Therefore, even though mesh results are better for the visualization since they lead to smoother shapes, in this context it is crucial to access all the information content, thus avoiding approximations as much as possible.

Gray level morphological operations One of the most relevant aspects of the proposed approach is the emphasis on objective evaluation. The influence of human factors in segmenting can be reduced also via pre-processing operations applied to the original MRI volume. This procedure is of fundamental importance when dealing with low-intensity field tomography, where it may be difficult to have accurate ground-truth. The proposed pre-processing operations are gray-levels morphological dilation and erosion. These techniques are well-known and widely described in the literature [183], therefore, details on their use are provided as dealing with the effects on the present segmentation evaluation process. Since, according to the structuring element, dilation and erosion take into account the information related to a voxel neighborhood, their application turns in reducing small differences between segmentation results when including or excluding a limited set of points. As a result, the esti-

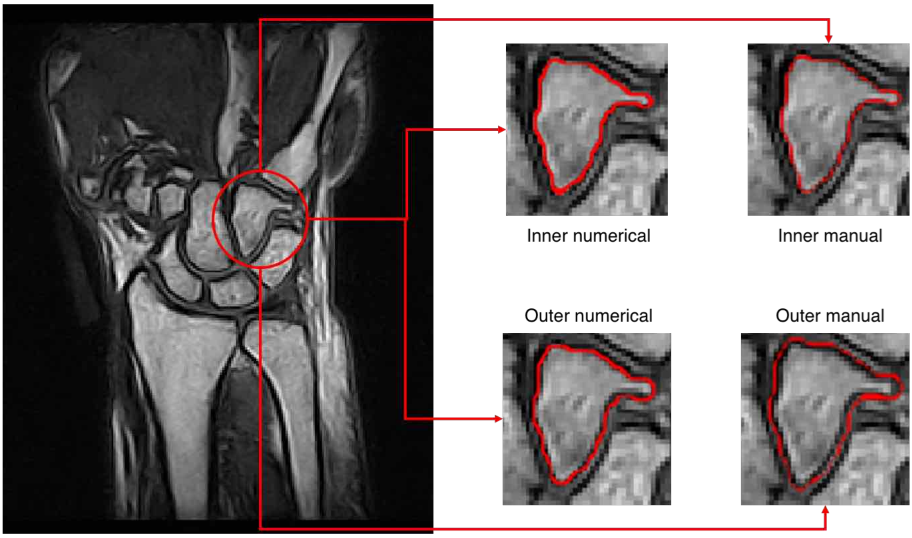


Figure 5.8: Example of the extracted boundary from the available segmentation.

mated gradient measure benefits of a better adaptation to the actual separation between cortical and trabecular bones, i.e., between bright and dark regions.

(a) Original (b) Estimated gradient (c) Dilation (d) Erosion

Figure 5.9: Original slice (a) and gradient estimation (b) as the difference between dilation (c) and erosion (d).

Boundary gradient estimation The integration between border and volume information is performed during the boundary gradient estimation phase. The estimated gradient in each boundary voxel results in the difference between the corresponding dilated and eroded gray levels. Such an operation is performed for both

BORDER$_1$ and BORDER$_2$.

Fuzzy criterion For defining the comparing procedure, let $I = [a, b]$ be the interval of feasible gradient values and, without loss of generality, let such an interval be of the form $J = [0, m]$ via the mapping $I \ni x \mapsto x - a \in J$ (thus $m = b - a$). Then, let $g_i \in J$ be the gradient value corresponding to the i^{th} voxel and let μ be the linear fuzzy membership function defining the punctual goodness of segmentation, namely how large is the gradient value in each voxel, with respect to the maximum: $\mu(i) = g_i/m \ \in [0, 1]$. The better the segmentation in the i^{th} voxel, the closer to 1 is $\mu(i)$. When comparing two segmentation results, the relation between them is based on order statistics. In particular, the minimum, median and maximum values of μ are considered; then, a voting procedure is defined. The rationale is that the segmentation that better fits the cortical-trabecular interface is the one with at least two highest values between the minima, medians, and maxima of μ.

Experimental results

The validation phase was conducted on a set of 98 3D T1-weighted low-field MR volumes of the hand-wrist district, which have been acquired by a 0.2 Tesla Artoscan (Esaote S.p.A., Italy), from the database in [7]. The size of each slice in the coronal plane is 256×256, while moving in the longitudinal direction, slices are between 90 and 120. Boundary voxels were obtained from both the segmentation of medical operators, and via the numerical method in [24], following the aforementioned procedure.

Statistical reasoning A globally good result should be characterized by large gradient values along with small occurrences of low gradient values. A possible local segmentation error is associated with low gradient values, which indeed represent minor variation between inner and outer gray levels. In order to test whether a segmentation result has local errors, it is mandatory to analyze its smallest histogram percentiles. As a consequence, such an analysis, as well as the comparison between two segmentation results, cannot be based on global statistical measures but must rely on local statistics. At the same time, the debate about central tendency measures is still open when dealing with their ability in catching useful and reliable information. Furthermore, the smallest gradient values interval cannot be uniquely defined since it depends on the ROI, on the specific tomographic device used, and is not normalized in MRI acquisitions. Indeed, the comparison of segmentation results is conducted via the analysis on the cumulative gradient histogram (CGH), which maps the various percentile values. More specifically, the comparison between two segmentation results is based on the low percentiles of the CGH. The best fitting to the cortical-trabecular interface is related to the CGH whose low percentiles correspond to higher values, meaning that few small gradient boundary points are present. Gradient values corresponding to CGH percentiles (10^{th} and 25^{th}) are summarized in Table5.3.

Table 5.3: Average results.

	Inner Numerical	Inner Manual
$10^{th}percentile$	516.35 ± 155.82	723.72 ± 120.46
$25^{th}percentile$	792.21 ± 101.35	978.04 ± 59.92
	Outer Numerical	Outer Manual
$10^{th}percentile$	350.91 ± 31.94	340.02 ± 36.73
$25^{th}percentile$	507.77 ± 96.69	481.67 ± 145.70

The boundary BORDER$_1$ yields to focus on the interior voxels, thus balancing the estimation error performed by the human user. On the contrary, the correct position of numerical segmentation moves the gradient evaluation inside the trabecular bone, thus yielding lower gradient values. Conversely, the boundary BORDER$_2$ delivers the highest ratio between the gradient from the unsupervised segmentation and the gradient from the manual one. Such values mean that many voxels on the border have low gradient in the manual approach, thus confirming the assumption that the medical operators included not only bright voxels, but also dark ones, as a consequence of their knowledge/expertise. On the other hand, the high border gradient for the unsupervised approach shows that such a method better fits the cortical-trabecular interface, thus clearly separating bright and dark voxels.

Case study The comparison between the proposed approach and the literature was performed on six volumes from the available database, which refer to three patients who have been evaluated twice by the doctors (time T0 and time T1, after 12 months). In particular, cases 1 and 2 refer to the first patient, 3 and 4 to the second one, and 5 and 6 to the third one.

In general, according to the fuzzy criterion here described, when considering the boundary $BORDER_1$, the best fitting to the cortical-trabecular interface is delivered by the manual segmentation. Conversely, in the case of boundary $BORDER_2$, the numerical segmentation is declared to be the best to fit the interface and highlighting the local sensitivity of the proposed criterion. Tables 5.4 and 5.5 show the segmentation with majority of voting for each bone and for each case, with respect to $BORDER_1$ and $BORDER_2$ respectively.

Table 5.4: Best fitting with respect to $BORDER_1$.

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
Capitate	Manual	Manual	Manual	Manual	Manual	Manual
Hamate	Manual	Manual	Manual	Manual	Manual	Manual
Lunate	Manual	Manual	Manual	Manual	Manual	Manual
Scaphoid	Manual	Manual	Manual	Manual	Manual	Manual
Trapezium	Manual	Manual	Manual	Manual	Manual	Manual
Trapezoid	Manual	Manual	Manual	Manual	Numerical	Manual
Triquetrum	Manual	Manual	Manual	Manual	Manual	Manual
Pisiform	Manual	Numerical	Manual	Manual	Manual	Numerical

Table 5.5: Best fitting with respect to $BORDER_2$.

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
Capitate	Numerical	Numerical	Manual	Numerical	Numerical	Numerical
Hamate	Numerical	Numerical	Numerical	Manual	Numerical	Numerical
Lunate	Manual	Numerical	Numerical	Numerical	Manual	Numerical
Scaphoid	Numerical	Numerical	Numerical	Numerical	Numerical	Numerical
Trapezium	Numerical	Numerical	Numerical	Numerical	Numerical	Numerical
Trapezoid	Numerical	Numerical	Numerical	Numerical	Numerical	Numerical
Triquetrum	Numerical	Numerical	Numerical	Numerical	Numerical	Numerical
Pisiform	Numerical	Numerical	Numerical	Numerical	Numerical	Numerical

The reader may notice that when considering $BORDER_1$ the manual annotation in inaccurate for three cases, in particular twice for the Pisiform and once for the Trapezoid. This may be due to the small size of such bones which make the manual annotation hard to perform. Conversely, when considering $BORDER_2$ the Numerical annotation is less accurate in four cases, once for the Capitate and the Hamate, twice for the Lunate. Such bones are the biggest ones and also the most regular thus also the manual annotation is easier to be performed. It is worth noting that the manual annotations performed on the same patient but at different times may be so dissimilar that in some cases, the fuzzy criterion here described did not agree on the same bone at different times ($BORDER_1$: Case 1-2 Pisiform, Case 5-6, Trapezoid and Pisiform; $BORDER_2$: Case 1-2 Lunate, Case 3-4 Capitate and Hamate, Case 5-6 Lunate).

Table 5.6: Entropy differences between numerical and manual segmentations.

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**
Capitate	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-5}$	$<10^{-5}$
Hamate	$<10^{-4}$	$<10^{-5}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$
Lunate	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-4}$
Scaphoid	$<10^{-5}$	$<10^{-4}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-5}$
Trapezium	$<10^{-4}$	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
Trapezoid	$<10^{-4}$	$<10^{-4}$	$<10^{-5}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
Triquetrum	$<10^{-5}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$	$<10^{-4}$
Pisiform	$<10^{-3}$	$<10^{-4}$	$<10^{-4}$	$<10^{-3}$	$<10^{-5}$	$<10^{-4}$

When applying the entropy-based evaluation method [184], the differences between the two approaches for the segmentation are not visible.

In order to apply the method in [184], let R be a segmented region of cardinality $|R|$, f a feature, and $R(f)$ the set of all possible values associated with feature f in region R. Then, let m be a feasible value for the feature f and $L_R(m)$ the number of voxels in region R that have a value m for feature f (i.e., L_R is the histogram of the feature f). The entropy for region R is defined as

$$E = -\sum_{m \in R(f)} \frac{L_R(m)}{|R|} \log \frac{L_R(m)}{|R|}. \tag{5.15}$$

The entropy measures the uniformity of the segmented region, based on a particular feature which, here, is the gray level of each voxel. The smaller the entropy, the more uniform the region, thus, the better the segmentation. However, when computing both the entropy of the numerical E_{num} and manual E_{manual} segmentation, one can notice that they are not very different, in particular $|E_{num} - E_{manual}| \leq 10^{-4}$.

In addition, one can compute the expected region entropy when referring to the whole district, thus evaluating the inner uniformity of each segmented bone. In particular, let K be the number of segmented regions, R_k be the k^{th} region of cardinality $|R_k|$, $k = 1, \ldots, K$, and $E(R_k)$ its entropy. The expected region entropy is defined as

$$\bar{E} = \sum_{k=1}^{K} \frac{|R_k|}{|R_{tot}|} E(R_k), \tag{5.16}$$

where $|R_{tot}|$ is the total number of voxels considered. However, even in this case $|\bar{E}_{num} - \bar{E}_{manual}| \leq 10^{-4}$, thus denoting that such an evaluation method does not recognize any difference between the numerical and manual approaches.

Chapter 6

Application of telerehabilitation techniques

The present chapter is devoted to the presentation of the results from studies involving the ReMoVES system. First of all, a feasibility study referring to the work in [66] is described. Then, the applications to SSc, USN, and post-stroke and frailty rehabilitation are presented, referring to [68], [94, 95, 93], and [138, 185] respectively. To conclude, the first phase referring to an ongoing study on balance assessment is disclosed.

The translational approach of the work underlying the present thesis, which combines the background of both the Polytechnic and Medicine Schools, is particularly evident in the present chapter, and represents its innovative contribution. Indeed, the previously introduced technology of the ReMoVES system, is deployed into real-world cases,and by leveraging on the application of data and signals analysis techniques, it results in a support to the clinical staff.

6.1 SVM-based feasibility study

A feasibility study for the deployment of a method for the automated evaluation of performance in exergames for motor rehabilitation is here presented. The considered data come from different patients as such a study was perform to test the applicability of ReMoVES for remote monitoring, aimed at using it for disease-specific treatment.

The proposed method is based on an SVM classifier, as classification should be based on the feedback and experience of the therapist. Indeed, the training phase of such model was performed on data labeled by the therapist as *Good* or *Other*, while the patients are performing their activity. SVMs are commonly used for motion analysis and are designed for two-class classification problems such as the one discussed here.

The workflow described in the following Sections should be applied separately to each exergame, with appropriate adjustment due to the differences among them. Here, the exergames *Shelf Cans* is considered.

Aimed at increasing the granularity of data, each session was split into segments, which are the actual input to be fed to the model (Figure 6.1). The segment duration ranges between 2.0 and 20.0 seconds with 0.1 steps. As a result, up to 60 samples can be obtained from each session.

Figure 6.1: Sessions are divided into segments of duration t, keeping the corresponding label.

The SVM classifier with RBF kernel has been trained through an iterative process for each t. Hence, in the proposed workflow (Figure 6.2), t is an hyperparameter that must be tuned according to the result of a grid-search process.

In order to analyze each segment, the following indicators are considered as feature vector for the Shelf Cans exergame.

- Game score.
- Hand trajectory, i.e., the average distance between the position of the hand and the shortest line passing through the items origin and targets.
- Range of motion of the elbow, i.e., the difference between the maximum and minimum angle calculated with the Wrist, Elbow and Shoulder coordinates in 3D space.
- Lateral flexion of the trunk.

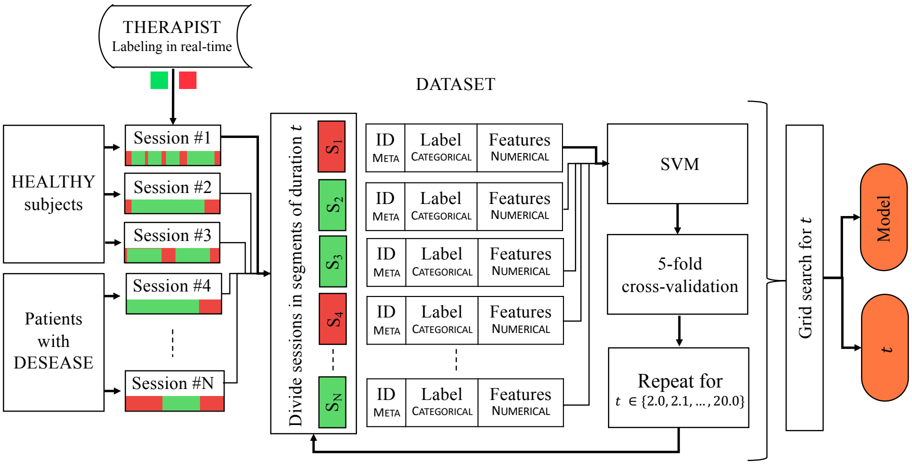


Figure 6.2: General flowchart of the SVM-based method.

- Trunk rotation on the transverse plane.

In the experimental phase, 30 full sessions of 120 seconds each were considered, resulting in hundreds of segments of duration t. The grid-search for the hyperparameter t is performed by training the corresponding models and evaluating their performance by iterating the 5-fold cross-validation for t between 2.0 and 20.0 seconds with 0.1 steps. The resulting accuracy for every t is taken into account. Non-linear least squares was used to fit the accuracy data and understand whether it presented a global maximum. As depicted in Figure 6.3, the fitted curve has a maximum in $t = 8.3$ seconds, thus being the chosen value for the hyperparameter t.

Then, the model was trained with samples generated by the duration segment $t = 8.3$, resulting in 389 samples in the dataset (200 for *Good* class and 189 for *Other* class) and the resulting overall accuracy of the model based on 5-fold cross-validation is 0.80 ± 0.16. The confusion matrix is provided in Tab. 6.1.

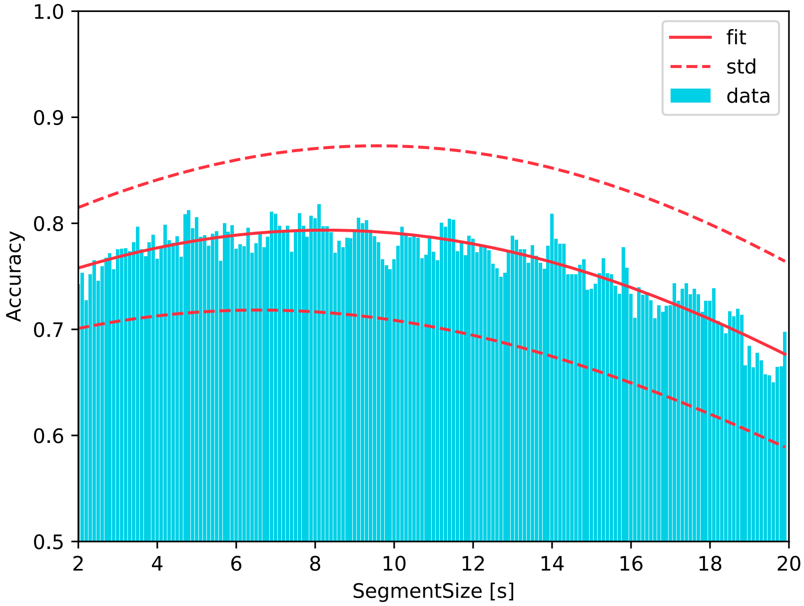


Figure 6.3: Plot of accuracy for each iteration of the grid-search procedure. The red curve fitting the data has the global maximum in $t = 8.3$.

Table 6.1: The final result of SVM performance expressed as a confusion matrix.

		Predicted		
		Good	*Other*	Total
Actual	*Good*	128	72	200
	Other	7	182	189
	Total	135	254	389

6.2 Application 1: Systemic sclerosis

The aim of the study[1] here presented is to assess the conditions of SSc patients, based on the activities delivered by ReMoVES . For such a purpose, it is necessary to verify that game sessions by patients and healthy subject are different. Therefore, the considered

---

[1] The study on SSc was conducted in collaboration with Struttura Complessa di Reumatologia - ASL3. In addition to the paper [68], it was also presented to the medical congresses ACR Convergence 2020 and 57° Congresso Nazionale SIR.

population was composed by both patients and healthy subjects (i.e., the clinical staff of the hosting facility).

The admission of patients to each game session was determined by the judgment of the clinical staff, who evaluated the willingness to participate and the general conditions of the patient at that particular time.

The treatment plan was divided on the basis of finger and wrist movements, in particular it consisted of:

- one session for stimulating thumb opposition (Finger Tap exergame);
- one session for stimulating finger flexion-extension (Floating Trap exergame);
- one session for stimulating radial-ulnar deviation (Endless Zig exergame);
- one session for stimulating wrist flexion-extension (City Car exergame).

All the games were performed with the dominant hand.

Patients feedback related to ReMoVES was extremely positive and all of them were willing to repeat the activities.

Table 6.2 summarizes the involved population with regard to each exergame.

Exergame	Group
Finger Tap	Patients: 19
	Control: 16
Floating Trap	Patients: 19
	Control: 16
Endless Zig	Patients: 18
	Control: 15
City Car	Patients: 18
	Control: 14

Table 6.2: Number of sessions for each exergame.

In order to compare the sessions performed by patients and control group, the Kruskal-Wallis test was used. Such a test focused only on one feature for each exergames, which follows.

- Finger Tap: number of correctly stopped marbles.
- Floating Trap: opening ROM, i.e., the difference between the maximum and minimum of Sphere field, which is the length of the radius of a virtual sphere contained in the hand palm (measuring the finger extension).
- Endless Zig: yaw ROM, i.e., the difference between the maximum and the minimum angle of the line joining the Wrist and the Palm positions on the coronal plane (measuring the radial-ulnar deviation).
- City Car: pitch ROM, i.e., the difference between the maximum and the minimum angle of the line joining the Wrist and the Palm positions on the sagittal plane (measuring the wrist flexion-extension).

The resulting p-values reported in Table 6.3 prove that the considered features are significantly different in the two groups.

Exergame	p-value
Finger Tap	0.0001
Floating Trap	0.0019
Endless Zig	0.0028
City Car	0.0001

Table 6.3: p-values from the Kruskall-Wallis tests.

In addition to the aforementioned analysis, an LSTM RNN was developed and tested, in order to provide an automatic evaluation of game sessions. Such an approach was applied only to Finger Tap, because of the nature of the movement involved. Indeed, thumb opposition is a small movement and hence, one can consider short sub-sessions which still depict appropriately the hand gesture.

Each Finger Tap session was divided into 5 segments. To avoid errors due to incorrect hand positioning either at the beginning or end, such starting and final frames were removed (1.5 sec at the beginning and 1.5 sec at the end). As a result, for each patient and subject, 5 segments collecting 114 extracted features each were fed in the network. Training, validation, and test sets were defined by splitting the data with percentage of 80%, 10%, and 10% respectively, hence enabling 10-fold cross-validation, in order to control for over-fitting. The network was trained from scratch. To account for randomness, cross-validation was performed on 15 permutations and the average overall accuracy was 86.31%.

6.3 Application 2: Unilateral spatial neglect

The ReMoVES activities used for assessing USN[2] disabilities are the digital versions of Albert's test, Line Bisection test, and Apple test. They are commonly applied in mapping visuo-spatial neglect in clinical practice. The system performs all the required data analysis to provide a scoring report to the medical expert. Additional information with respect to paper-and-pencil version is provided to offer a complete diagnosis. The feasibility of the method is based on a comparative analysis in which the parameters normally extracted by the therapist observing the test are compared with the parameters automatically derived from the digital version. The implemented test version is defined in the following, along with the data and key indicators that enable monitoring of the actual performance of patients. Then, an real-world application is disclosed with regards to two case-studies.

The digital version of the Albert's test can provide more granular information about the patient's exploration capability. Recall that in this test, patients are required use a pencil to cross out forty 2.5 cm lines which are positioned in pseudo-random orientations on a piece of paper. The actual disposition of these lines is standardized, allowing for a systematic analysis of subjects' performance on the left, on the right, and in the middle of the page. More specifically, in the Albert's test, the paper is divided in seven sections by column, which are hereinafter numbered from left to right. Scoring is based on the number and location of lines left uncrossed, in particular USN

---

[2] The studies on USN were conducted in collaboration with Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, and Maternal and Children's Sciences (DINOGMI) - Università degli Studi di Genova, Clinica di neuroriabilitazione - IRCCS Ospedale San Martino, and Struttura Complessa Recupero e Rieducazione Funzionale - ASL3. In addition to the papers [94, 95, 93], some results were also presented to the medical congress XX Congresso Nazionale SIRN.

is indicated when these lines are on the same side of the page as the patient motor deficit is located. This can be done by calculating the covered area index and the .

A first explanatory index that can be defined is the *covered area index*:

$$Covered\ Area\ Index = \frac{A_{explored}}{A_{max}}, \tag{6.1}$$

where $A_{explored}$ is the area of the smallest rectangle containing the crossed lines, and A_{max} is the area of the smallest rectangle containing all the lines.

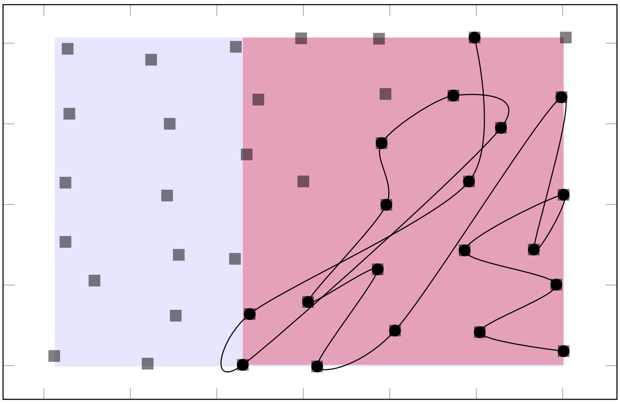


Figure 6.4: Red area represents the area explored. The total area containing all the lines is in blue color.

Another index is the center of gravity (CoG), which is a point whose coordinates are the mean value of x and y coordinates of the crossed out lines.

Three other indicators described in the following, can be collected from the assessment and the analysis of digital version of the test. The *execution order* is the order in which the patient has crossed out the lines; the *distance* between two crossed out consecutive targets; the *time* that elapses between the consecutive cancellation of two targets.

Then, the Line Bisection test is assessed by calculating the percent deviation (PD) of the bisection from the center of the line using the following formula [186]:

$$PD = \frac{(MeasuredLeftHalf - TrueHalf)}{TrueHalf} \times 100. \quad (6.2)$$

Recall that here, the patient must place a pencil mark at the center of a series of horizontal lines. Therefore, the resulting score is lower as the pencil mark is far from the center of the line. Therefore, the main indicator for Line Bisection test is the *offset*, i.e. the mean of the measured deviations during the whole session.

Finally, the activity Apple Test is the digital version of the Italian standardization of the Apples Cancellation Test [86]. Apple test is a cancellation task in which outline drawings of 150 apples are shown scattered over a sheet. Several apples are presented in an upright position. One-third of the apples are full (targets), and two-thirds are open on either the left or on the right side (distracting elements). In order to balance the probability that omissions will show left versus right or upper versus lower space neglect, the sheet is virtually divided into a grid with two rows and five columns to ensure an equal distribution of the apples across the page. Each cell of the grid contains 15 apples: three large ones (one with no opening, one with an opening on the left and one with an opening on the right) and 12 small apples (four without openings, four with openings on the left and four with openings on the right). The large apples are 50% bigger than the small ones. In the digital version, each participant is asked to touch on the screen all the full apples and to ignore all the ones with holes.

ReMoVES automatically calculates the score indices, namely *total number of crossed-out targets*, *asymmetry score for egocentric neglect*, and *asymmetry score for allocentric neglect*. The maximum value of the first index is fifty, that is the total number of full apples. The *asymmetry score for egocentric neglect* is the difference between the number of targets on the right side and the number of targets on the left side of the page, excluding the apples contained in the upper and the lower cells of the middle column. The maximum score is twenty. The *asymmetry score for allocentric neglect* is the difference between the number of targets crossed out with a left opening and the number of targets crossed out with a right opening. In both egocentric and allocentric neglect, negative asymmetry scores denote right neglect and positive asymmetry scores denote left neglect.

In a preliminary study twelve patients suffering from USN were involved, aimed at comparing the standard and digital version of the Albert's test. Ten of them presented right lesions, thus a left-sided neglect and two of them have left lesions, thus corresponding to a right-sided neglect. For a further comparison, a control group of six healthy subjects is also considered.

Data was collected during the first session, and all patients performed the paper-and-pencil test first and then the digital test. It is worthy to claim that the intervention using the ReMoVES exergames platform is not intended to replace the classic practice, but it is an integrative tool of usual treatment.

First of all, the correlation between standard and digital version of the Albert's test is computed. For each patient and for each version, the percentage of crossed out targets in each column is collected into a 1×7 vector, called p_t and p_d for the traditional and digital version respectively. Then, the correlation coefficient (Pearson's coefficient) between p_t and p_d is calculated. The resulting average correlation coefficient is $corr = 0.82$, thus denoting a good concurrent validity of the digital version with respect to the standard one. Figure 6.5 depicts the crossed out target percentages in both the traditional (blue) and digital (red) version of the test. In general, blue and red crosses are very close for each patient, thus confirming the assumption of good correspondence between the traditional and digital versions. Taking a deeper look to the graph one can notice that patient 5 erases more targets in the digital test, while patient 6 performs better in the paper test. This is probably due to the difficulty in holding a pen and in using the touchscreen they have respectively.

In order to investigate the sessions based on the ignored hemispace, in Table 6.4 only the ten patients with left neglect are considered. The average percentages of crossed target are lower in the first sections, and larger in the last ones, thus confirming the left-sided neglect. Furthermore, one can also notice the similarities between the percentages of the two versions of the test.

In order to assess the homogeneity of the available population, an

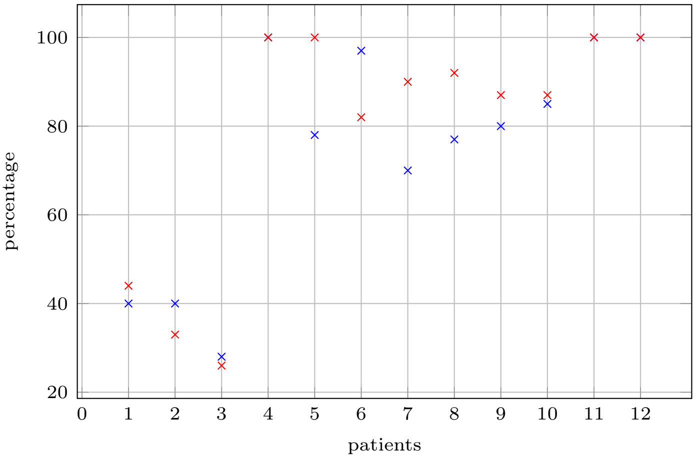


Figure 6.5: Crossed out target percentages, blue is for the standard version of the test, red is for the digital one.

	Traditional test	**Digital test**
Section 1	42%	58%
Section 2	62%	65%
Section 3	68%	70%
Section 4	72%	82%
Section 5	78%	92%
Section 6	90%	92%
Section 7	98%	96%

Table 6.4: Average percentages of targets erased in the seven sections, in both traditional and digital version of Albert's test.

Analysis of Variance (ANOVA) test is performed on erased targets. The result of the ANOVA is the rejection of the null hypothesis and this denotes inhomogeneity in the population for both the results of the digital and traditional tests. Actually, it is an expected result because the population is composed by patients with both severe and minor injuries. For this reason, the population is split in two different groups according to the criteria suggested in [187]. In the contralesional part, a rate of more than 40% of omission was considered an index of severe neglect. Conversely, a rate of less than 40% of omission was taken as an index of mild/moderate form of USN. Respectively, the first group A consists of three patients and the others belong to group B. As a counterproof, two ANOVA tests are performed for each group formed and the results are the acceptance of null hypothesis (with $\alpha = 0.05$), meaning that the two groups are homogeneous, based on the results of the digital and traditional tests.

To sum-up, both the correlation study and the outcomes from the ANOVA tests prove that traditional and digital versions are interchangeable for the population under analysis.

In addition to the crossed out targets-related information, that can also be obtained from the paper version of the test, other automatic indicators that are computed by ReMoVES are here presented.

The trajectories followed by two patients are compared with the one followed by a healthy subject. In Figures 6.6, 6.7, and 6.8, three types of point are highlighted: the blue ones indicate the start; the red ones indicate the final point; and the green points indicate the CoG. It is worth mentioning that the CoG position is in agreement with the left- or right- sided neglect diagnosis, for each patient. Furthermore, the plots provide a feedback of the different strategies adopted to crossed out the targets. In general, healthy subjects performed regular paths, i.e. from left to right or from top to bottom. Figure 6.7 represents the trajectory performed by a patient belonging to group B, that denotes a more regular path compared to the path tracked by a patient of group A (Figure 6.6. Moreover, both patients have a left-sided neglect, in fact, the right side of the game area is the one from which patients started the test and that have explored more in depth.

From the digital version one can also acquire information related to the times between two target cancellations and the distance between two consecutive crossed out targets. For what concerns the average time, it is 3.63 sec for patients in group A, 2.00 sec for group B, and 0.51 sec for healthy subjects. As one could suppose, the average time directly depends on the disease severity, thus it can be an informative indicator for the USN assessment. Similarly, the average distance between two crossed targets is 3.09 cm for patients in group A, 2.81 cm for group B, and 2.33 cm for healthy subjects. The trajectory, and hence the cancellation strategy adopted by patients

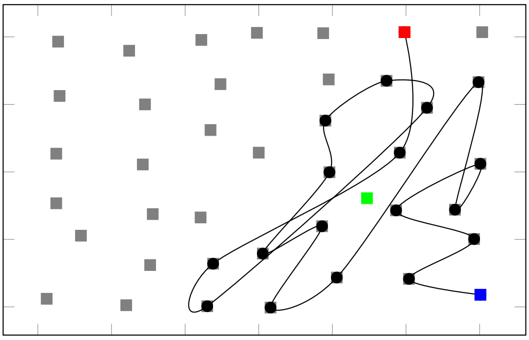


Figure 6.6: Trajectory followed by a subject in group A. The blue point is the starting point, the red point is the final one, and the green point is the CoG.

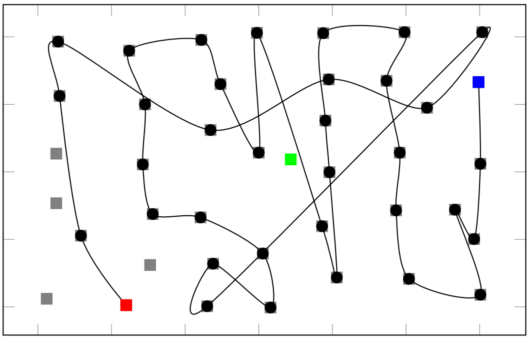


Figure 6.7: Trajectory followed by a subject in group B. The blue point is the starting point, the red point is the final one, and the green point is the CoG.

in group A seems pretty random, and that is reflected on the average distance extracted from their sessions. Conversely, group B follows a more regular path, which is denoted by the smaller distance between consecutive targets, and so by the average time value which is close to the one by the healthy subjects. Recall that high values of distances or times can also mean that patients have explored some part of the ignored area, where indeed they had difficulty in finding targets.

As anticipate in Section 2.3, ReMoVES exergames can be deployed as complementary cognitive therapy associated to the tDCS

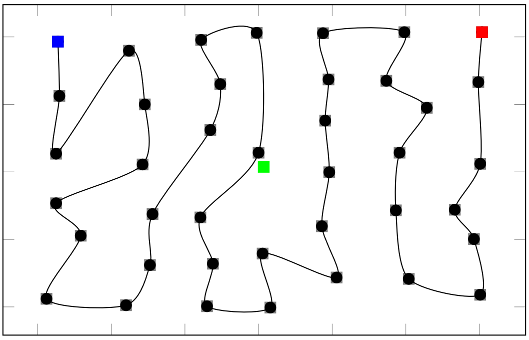


Figure 6.8: Trajectory followed by a healthy subject. The blue point is the starting point, the red point is the final one, and the green point is the CoG.

approach. In such a context, a real-world application is disclosed with regards to two case-studies. Both patients underwent ten A-tDCS consecutive sessions (five days per week over a two weeks period) coupled with simultaneous digital treatment (ReMoVES). During the training task A-tDCS (current of 1.5 mA) was delivered for twenty minutes. Patients tolerated the application of tDCS without any adverse side-effects.

Patient A is a 70 years-old right-handed woman, who started complaining with vague difficulty in reading, writing on the computer, driving and parking her car. She also experienced problems in finding objects on the left (e.g., in the kitchen while cooking). Entering the car from her left side (as a passenger) was almost impossible, while the same problem did not happen when getting in the car as a driver (right side). The patient was first seen by an ophthalmologist and then by a neurologist. Apart from a minimal bilateral myopia, the physical examination was normal. Optical Coherence Tomography (OCT), fluorescein angiography, visual evoked potentials and lectroretinogram were unremarkable. Automatic perimetry did not show hemianopia or other visual field defects. The neuropsychological examination showed the following findings: signs of simultanagnosia, writing and reading disturbances, and left neglect. The Mini Mental State Examination (MMSE) was 25/30 (cut-off ≥ 24). the patient showed severe impairment when reading, writ-

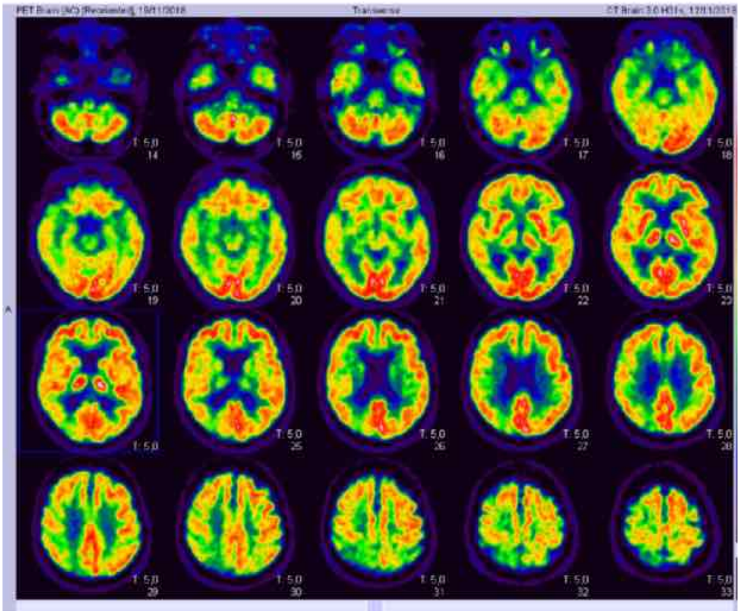


Figure 6.9: Patient A: cerebral PET showing a significant hypometabolism in the posterior temporal and occipital cortex and in the bilateral parietal cortex.

ing and copying (intersecting pentagons), while the verbal memory, the phonemic and semantic fluencies were intact. Cerebral MRI revealed an atrophy in posterior temporal and occipital cortex and in bilateral parietal cortex (right worse than left), and the PET showed a significant hypometabolism in the same areas (Figure 6.9).

A diagnosis of Posterior cortical atrophy (PCA) was made. PCA is a neurodegenerative condition characterized by an insidious onset, gradual progression and prominent early disturbance of visual functions, in the absence of ophthalmologic causes [188]. The most frequently neuropsychological deficits are: alexia and features of Balint's syndrome (simultanagnosia, oculomotor apraxia, optic ataxia, environmental agnosia) and Gerstmann's syndrome (acalculia, agraphia, finger agnosia, left/right disorientation) and neglect [189].

After the tDCS treatment, Patient A showed a clear improvement in the line bisection test (Mann-Whitney U test, $p < 0.003$), a less pronounced amelioration in the cancellation tests, while the drawing test was still impaired.

The improvement of the exploration capability of Patient A is monitored via the *covered area index*. Figure 6.10 shows the covered area index variation along the sessions.

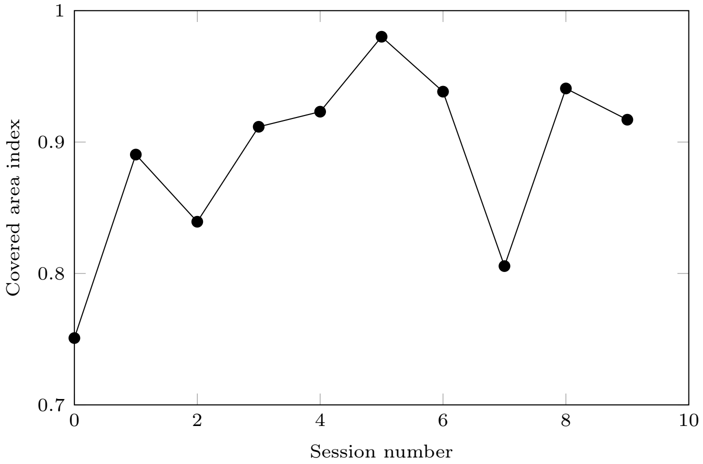


Figure 6.10: Patient A: covered area index variation along the sessions. No regular behaviour can be noticed, thus proving the absence of learning effect on the task.

The path followed by the patient to cancel the targets provides a feedback on the strategy adopted to fulfill the task. Figures 6.11 and 6.12 depict the trajectories at time $T0$ and $T1$. The more regular path at time $T1$ shows a clear improvement in the patient's conditions, thus denoting a rehabilitative success.

The adopted strategy can also be evaluated by calculating the distance between two consecutive canceled targets: the smaller the distance, the more efficient the strategy. Figures 6.13 and 6.14 show the histogram of approximate distances between consecutive canceled targets at time $T0$ and $T1$ respectively. The number of occurrences on the bars on the left is larger at time $T1$ than the one at time $T0$, thus showing a better strategy adopted after the rehabilitation treatment, coherently with what can be inferred from the trajectory analysis. The average distance at time $T0$ is 3.71 ± 1.86 cm, the average distance along the rehabilitation process is 3.17 ± 1.83 cm, and the average distance at time $T1$ is 3.02 ± 1.82 cm (the average

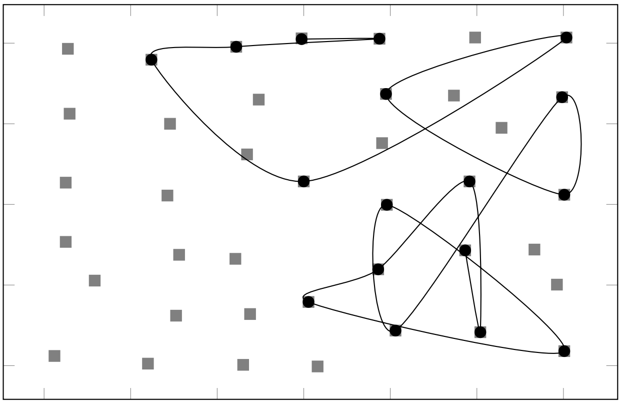


Figure 6.11: Patient A: trajectory reproducing the order of targets cancellation at time $T0$.

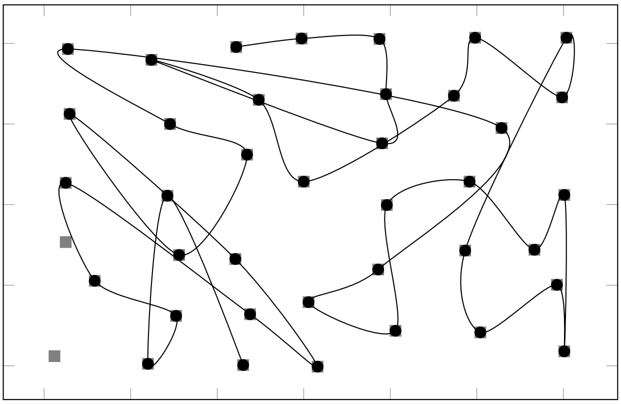


Figure 6.12: Patient A: trajectory reproducing the order of targets cancellation at time $T1$.

distance of a healthy subject 2.38 ± 1.26 *cm*).

To conclude, ReMoVES platform provides also the waiting time between two consecutive cancellations. A small time between targets cancellations denotes that a good strategy has been adopted. Conversely, a large waiting time means that the patient is exploring a new region of the game area, thus requiring more time to detect a new target. Figure 6.15 and 6.16 show the variation of waiting time between two consecutive cancellation at time $T0$ and $T1$. The

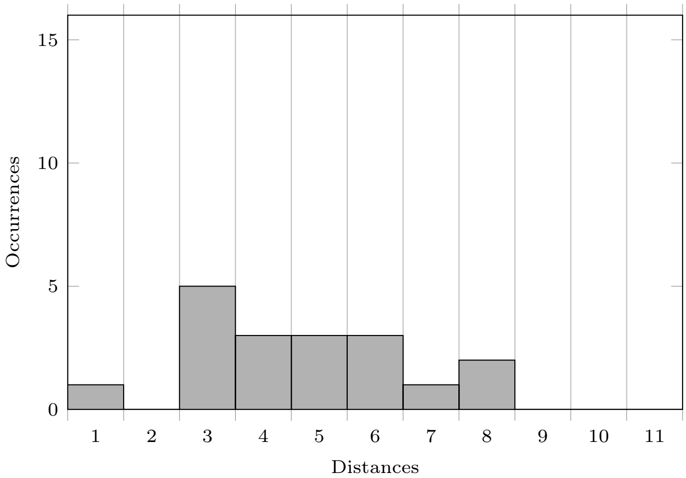


Figure 6.13: Patient A: histograms of approximate distances between a canceled out target and the following one at time $T0$.

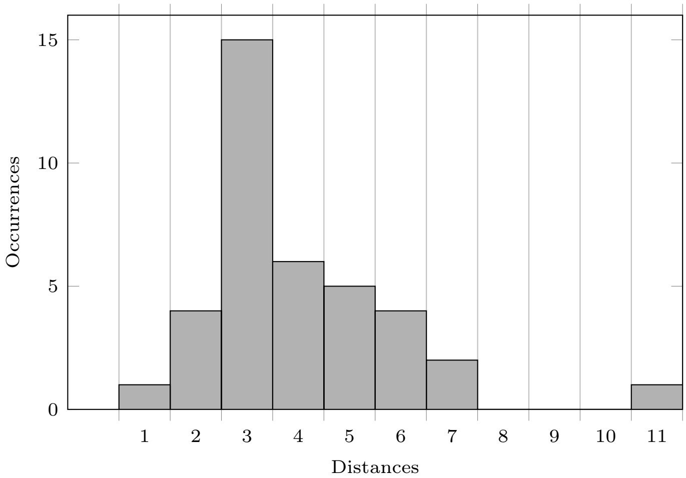


Figure 6.14: Patient A: histograms of approximate distances between a canceled out target and the following one at time $T1$.

waiting time increases at the end of the session, until the patient states that non more targets are present. The average waiting time at time $T0$ is 49.11 ± 29.15 *sec*, the average waiting time during the rehabilitation process is 34.70 ± 27.43 *sec*, and the average waiting time at time $T1$ is 27.50 ± 21.78 *sec* (the average waiting time of a healthy subject is 5.95 ± 1.64 *sec*).

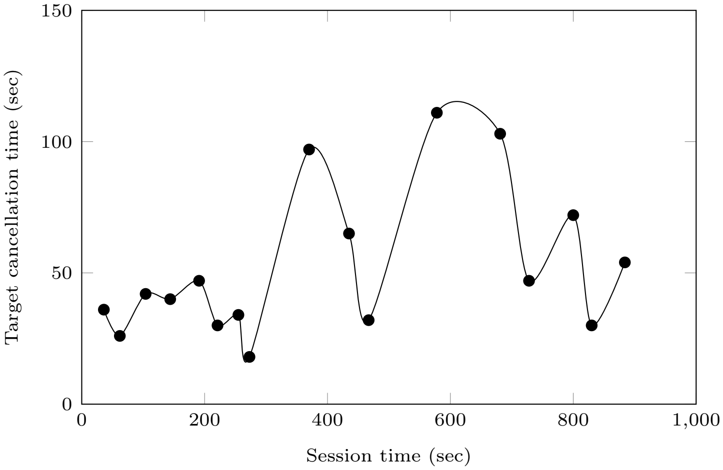


Figure 6.15: Patient A: curves depicting the variation of waiting time between a cancellation and the following one at time $T0$.

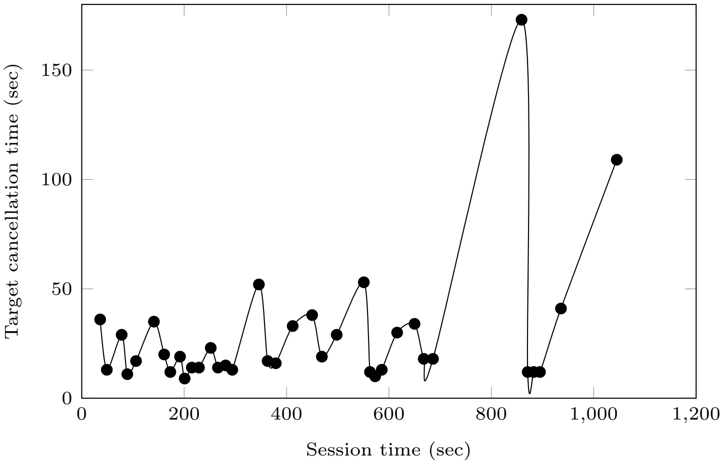


Figure 6.16: Patient A: curves depicting the variation of waiting time between a cancellation and the following one at time $T1$.

Patient B is a 68 years-old man who was referred to the Neurological Rehabilitation Unit of Ospedale Policlinico San Martino IRCCS of Genoa with sequelae of a right cerebral hemorrhage. CT revealed wide right nucleoscapular intraparenchymal hematoma in the homolateral temporal-frontal parietal area (Figure 6.17).

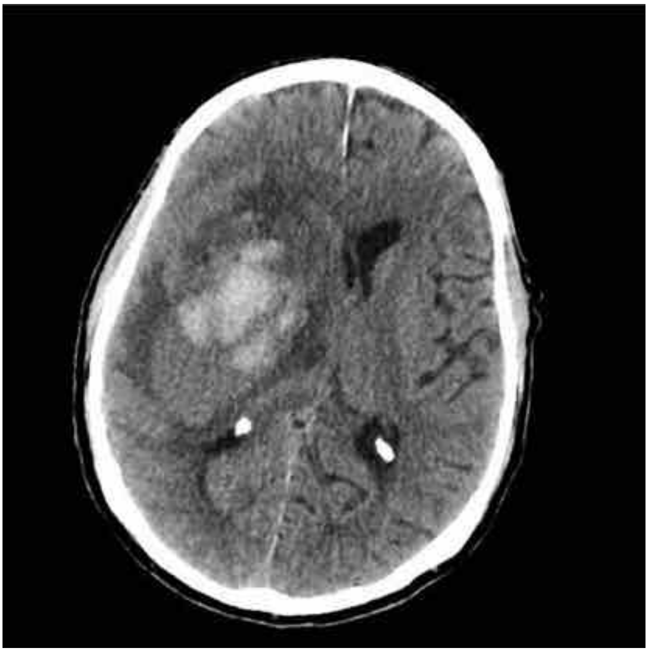


Figure 6.17: Patient B: cerebral CT shows a large right intraparenchimal hemorrhage with surrounding edema involving the fronto-temporo-parietal areas. Recall that in the CT scans, the right hemisphere is shown on the left.

The physical examination showed left hemiparesis, normally fluent speech, impairment of superficial tactile sensitivity and proprioception on the left. In a sitting position, tendency to retroplusion and lateropulsion of the trunk towards the left. The neuropsychological examination showed left neglect while the other cognitive functions (i.e., orientation, memory, and language) were intact. After the tDCS treatment, the patient improved up to normal performance in the line bisection test, in the line cancellation test, and in the Apples test (Mann-Whitney U test, $p < 0.02$, $p < 0.01$, and $p < 0.05$ respectively). The drawing test was substantially normal but for an omitted detail on the left side.

Here, the game area is divided into ten sub-regions, accordingly to [86]. The average difference between both correctly and erroneously canceled targets in digital and paper versions is used as similarity metrics at time $T0$ and $T1$. In particular, at time $T0$ the average difference between correctly canceled targets is 1.6 globally, 0.75 in the four left sub-regions, and 1.75 in the four right sub-regions. Furthermore, at time $T1$ it is 1.1 globally, 2 in the four left sub-regions, and 0.25 in the four right sub-regions. As for the average difference between the number of canceled distractors, at time $T0$ it is 0.9 globally, 0.25 in the four left sub-regions, and 0.5 in the four right sub-regions, while, at time $T1$ it is 0.6 globally, 0.75 in the four left sub-regions, and 0.25 in the four right sub-regions. The average difference between the number of canceled targets in the two versions is smaller than 2, thus denoting a good similarity between digital and paper versions.

The ReMoVES platform provides the automatic calculation of the score indices suggested in [86], namely *total number of crossed-out targets*, *asymmetry score for egocentric neglect*, and *asymmetry score for allocentric neglect*. At time $T0$ they are respectively 25, 13, and 17, while at time $T1$ they are respectively 29, 8, and 5. The positive values of the asymmetry score for both egocentric and allocentric neglect at time $T0$ are coherent with the diagnosis of left neglect. Furthermore, the increase in the total number of crossed out targets and the decrease of the asymmetry score for both egocentric and allocentric neglect at time $T1$ denote an improvement of the patient's condition, and thus a rehabilitative success.

Likewise the Albert's test, the waiting time between two consecutive cancellations is provided. Figure 6.18 and 6.19 show the variation of waiting time between two consecutive cancellations at time $T0$ and $T1$. The average waiting time at time $T0$ is 56.56 ± 54.77 *sec*, the average waiting time during the rehabilitation process is 46.47 ± 44.40 *sec*, and the average waiting time at time $T1$ is 40.67 ± 32.59 *sec* (the average waiting time of a healthy subject is 8.29 ± 6.64 *sec*).

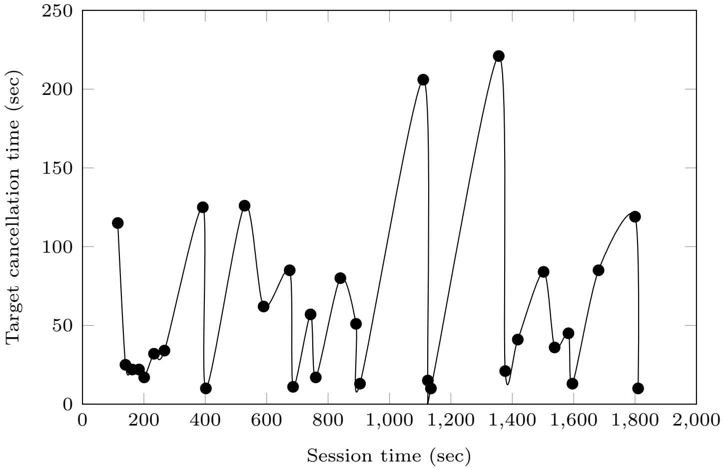


Figure 6.18: Patient B: curves depicting the variation of waiting time between the cancellation of a correct target and the following one at time $T0$.

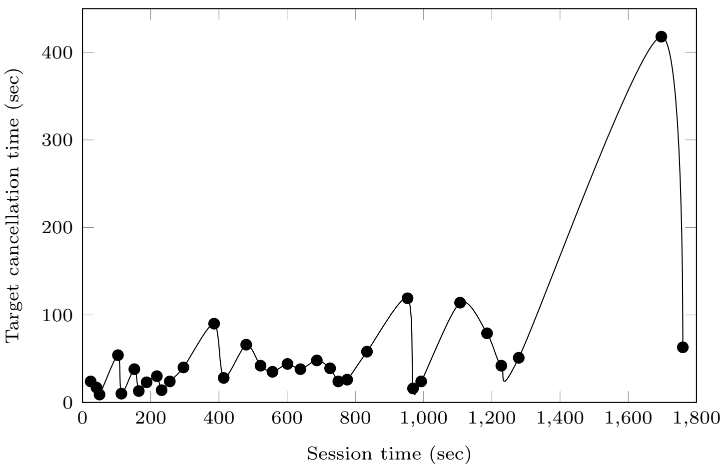


Figure 6.19: Patient B: curves depicting the variation of waiting time between the cancellation of a correct target and the following one at time $T1$.

6.4 Application 3: Post stroke

The application of ReMoVES for post stroke treatment[3] refers to the practice of the Shelf Cans exergame for stimulating the upper-limb movement.

The involved features enable clear visualizations, allowing clinicians to easily picture patients' performance, even without directly attending them. The experimental phase is conducted on two case-studies with regards to the evaluation of upper-limbs mobility in patients with ataxia. The adoption of the strategy of weighting the limb when performing the movement, i.e., placing some weight on the limb being involved in the movement, is discussed (see Section 2.3). The indicators computed in both sessions with and without strategy are compared, also referring to the practice of some healthy subjects. The current study presents a starting point for several possible applications: i) a support for a quantitative evaluation of therapy; ii) an easy way to remote control of training; iii) the definition of criteria for evaluating therapeutic strategy.

Patient A was a 49-year-old man, who suddenly experienced unsteadiness of gait, incoordinate movements of the right limbs, blurred vision and diplopia. The neurological examination showed a severe right ataxia and a bilateral gaze evoked horizontal nystagmus. Standing and gait were ataxic and broad-based. Brain MRI revealed a right cerebellar infarction. Two weeks later, the cerebellar signs improved moderately.

Patient B was a 59-year-old man who presented with sudden headache, nausea, speech disturbances and unsteadiness of standing and walking. The neurological examination revealed a moderate axial and right limb ataxia, together with slurry speech. Brain MRI showed a right cerebellar hemorrhage. After ten days, a slight improvement of the cerebellar signs and symptoms was observed.

After the acute phase of the disease, both patients were admitted

---

[3] The studies on stroke consequences were conducted in collaboration with Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, and Maternal and Children's Sciences (DINOGMI) - Università degli Studi di Genova, Clinica di neuroriabilitazione - IRCCS Ospedale San Martino, Centro di riabilitazione srl and Struttura Complessa Recupero e Rieducazione Funzionale - ASL3. In addition to the paper [138], some results were also presented to the medical congresses XVIII, XIX, and XXI Congresso Nazionale della Società Italiana per lo Studio dello Stroke.

to the Neurological Rehabilitation Unit of Ospedale Policlinico San Martino IRCCS of Genoa. The Scale for the Assessment and Rating of Ataxia (SARA) [190] was utilized to evaluate the severity of the cerebellar disorder. SARA values range from 0/40 (no ataxia) to 40/40 (most severe ataxia). Patient A scored 24/40 and Patient B scored 13/40. The Activity-Specific Balance Confidence (ABC) Scale [191] was also employed to assess the subjective confidence of balance (0% not safe at all, 100% completely safe). Patient A had 60%, while Patient B reported 74,3%. In conclusion, it is worth underlining that the two patients were similar in height and weight.

Patients data have been analyzed via a direct and an indirect approach. Direct analysis employs 3D coordinates of joints provided by Kinect sensor. According to the biomechanical model, the human body can be considered as a system of rigid bodies, i.e., a set of body segments connecting the joints. Each pair of adjacent segments is considered as a kinematic and rotational pair. Consequently, significant features can be extracted to describe patient's movements. Referring to the present exergame, the analysis is focused on the upper-limb movement. The patient is required to extend the elbow and move the shoulder in order to store the can, without tilting the trunk to compensate the lack of reserve movement.

The computed indicators are the range of motion of the shoulder and elbow in the coronal plane, and the range of motion of the trunk in the sagittal plane. In particular, let us consider a fixed time, and define the shoulder and elbow angles in the coronal plane, and the trunk angle in the sagittal plane as

$$\theta_{shoulder} = \arctan \frac{z_8 - z_5}{x_8 - x_5}, \tag{6.3}$$

$$\theta_{elbow} = \theta_{shoulder} - \arctan \frac{z_8 - z_{16}}{x_8 - x_{16}}, \tag{6.4}$$

$$\theta_{trunk} = \arctan \frac{z_4 - z_7}{y_4 - y_7}, \tag{6.5}$$

respectively. Notice that shoulder ad elbow angles have been defined with respect to the right arm; the definition for the left arm is straightforward. Hence, the range of motion of the shoulder and elbow in the coronal plane, and the range of motion of the trunk in the sagittal plane are defined as

$$ROM(\theta_{shoulder}) = \max \theta_{shoulder} - \min \theta_{shoulder}, \tag{6.6}$$

$$ROM(\theta_{elbow}) = \max \theta_{elbow} - \min \theta_{elbow}, \tag{6.7}$$

$$ROM(\theta_{trunk}) = \max \theta_{trunk} - \min \theta_{trunk}, \tag{6.8}$$

respectively. The coordinates involved in Equations 6.3, 6.4, and 6.5 refer to the spatial coordinates of joints in Figure 4.2, where x, y, and z represent the mediolateral, anteroposterior, and vertical directions, respectively.

For what concerns the indirect analysis, it was performed through 2D data obtained by the videogame. The three specific tasks of Shelf Cans, referring to the differently colored cans, are split. Then, the straight line connecting starting position and targets is computed, and will be hereinafter referred to as *the optimal trajectory*. In addition, the so-called *approximate trajectory* performed by the patient during the game session is computed as the regression line of hand-game positions during the considered task. The lower the angle between the two lines, the better and controlled movement was done by the patient. Indeed, small angles shows that the approximate fitted path is similar to the optimal one. Conversely, large angles are typical of trajectories which are far from the optimal one. The linearity of scope-oriented movement is usually valued during physical therapy for patients affected by pathologies of motor learning such as cerebellar stroke. The importance of such an indirect analysis relies on the fact that it allows for quantifying the degree of improvement of the pathology which has caused the deficit of movement, and also to quantify the motor learning.

For each patient, two sessions of Shelf Cans exergame were considered, namely with and without weighting the involved upper-limb. Indeed, according to Bhanpuri et al. [102] a cerebellar damage likely causes an inertial mismatch between an internal representation of body dynamics and the actual body dynamics. On this base, a hypometric and a hypermetric cerebellar patient would respectively underestimate or overestimate their limb's inertia. Adding mass to the affected limbs can have beneficial effect on such a mismatch. Some other authors, however, failed to replicate the beneficial effects of such a strategy, especially for multi-jointed reaching movements [103]. The patient played the exergame while sitting, in order to reduce trunk and arms oscillation in the standing position, which could negatively affect the collected data.

Here, the effect of weighting the limb is quantified in order to provide the clinical staff with such objective data. The weight amounts of one kilo, and was placed on the wrist of the used limb. The Microsoft Kinect sensor was place in front of the players, at a distance of three meters. The duration of the game session is 90 seconds. In order to avoid the influence of the order in using or not the weight for the activities, two hours time distanced the first and second session, so that they can be considered independent. In addition, some training session had been performed the days before these trials, for familiarizing with the system.

Values of the indicators extracted from both the sessions with and without the strategy are summarized in Tables 6.5 and 6.6, and the graphs referring to the indirect analysis can be visualized in Figures 6.20, 6.21. In Figure 6.23 values of the ROM are depicted.

Table 6.5: Angles between the approximate and optimal trajectories.

	Red can	Orange can	Green can
Pat. A no weight	7.18	63.77	13.96
Pat. A weight	6.25	5.28	2.32
Pat. B no weight	4.70	49.37	11.49
Pat. B weight	5.82	1.89	12.80

Table 6.6: ROMs of shoulder, elbow and trunk.

	Shoulder ROM	Elbow ROM	Trunk ROM
Pat. A no weight	14.66	54.86	17.01
Pat. A weight	149.47	108.25	7.72
Pat. B no weight	128.56	24.54	5.47
Pat. B weight	132.03	3.76	3.39

The weighting strategy yielded a more precise movement in the performance of Patient A, which is denoted by the decreased angle between the estimated and optimal trajectories, and also by the increased range of motion of both shoulder and elbow. In addition, the reader can notice that the range of motion of the trunk decreases when patient A has a weight on the limb. This is probably due to a major control on its arm that is enabled by the strategy.

The same strategy resulted less efficient for Patient B, where, apart from the angle between the estimated and optimal trajectories when handling the orange can, no other significant better performance could be detected. In general, indicators values are better for patient B, and his better general condition is likely the reason for a less visible effect of the strategy on the performance.

Aimed at favoring comparisons with healthy subjects, a control group of six persons was considered and took part to the same treatment as patients. The values of the indicators extracted from both the sessions with and without the weight are summarized in Tables 6.7 and 6.8. Also, an example of graph about the indirect analysis is depicted in Figure 6.22. In general, it is worth noting how the adoption of the weight strategy did not yield to relevant difference for healthy subjects.

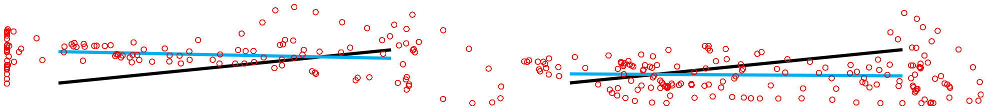
Table 6.7: Angles between the approximate and optimal trajectories in the control group.

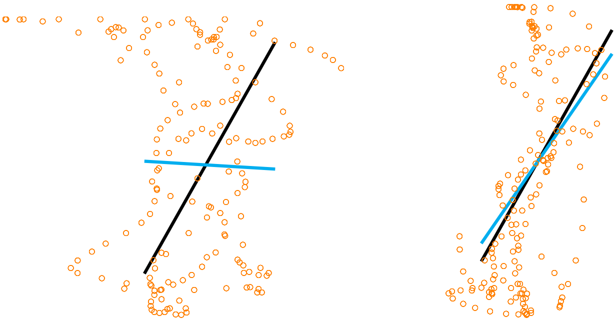
	Red can	Orange can	Green can
Sub. 1 no weight	6.29	1.75	4.55
Sub. 1 weight	5.02	2.02	5.69
Sub. 2 no weight	4.25	19.11	3.13
Sub. 2 weight	5.85	15.65	1.48
Sub. 3 no weight	4.21	28.26	1.15
Sub. 3 weight	6.57	13.29	2.69
Sub. 4 no weight	7.52	0.54	2.98
Sub. 4 weight	2.60	2.87	0.16
Sub. 5 no weight	0.42	0.35	0.92
Sub. 5 weight	1.24	1.03	0.60
Sub. 6 no weight	5.20	9.23	2.22
Sub. 6 weight	6.01	3.33	5.01
Avg. healthy no weight	4.65	9.87	2.49
Avg. healthy weight	4.55	6.37	2.61

Table 6.8: ROMs of shoulder, elbow and trunk in the control group.

	Shoulder ROM	Elbow ROM	Trunk ROM
Sub. 1 no weight	179.71	97.76	3.68
Sub. 1 weight	179.21	140.48	4.89
Sub. 2 no weight	179.58	168.21	10.64
Sub. 2 weight	178.72	139.03	12.73
Sub. 3 no weight	179.22	72.48	5.54
Sub. 3 weight	179.74	92.56	5.23
Sub. 4 no weight	179.43	68.34	3.71
Sub. 4 weight	179.95	36.12	3.18
Sub. 5 no weight	179.49	46.30	3.75
Sub. 5 weight	178.45	45.65	3.22
Sub. 6 no weight	179.80	79.77	16.33
Sub. 6 weight	178.77	74.65	12.70
Avg. healthy no weight	179.54	88.81	7.28
Avg. healthy weight	179.14	88.08	6.99

In general, this analysis is conducted to be delivered to the clinical staff, in order to help them in defining a personalized plan of care, and also to support patients in acquiring or reacquiring faculties to employ in daily life activities.


(a) Red can - no strategy (b) Red can - strategy


(c) Orange can - no strategy (d) Orange can - strategy

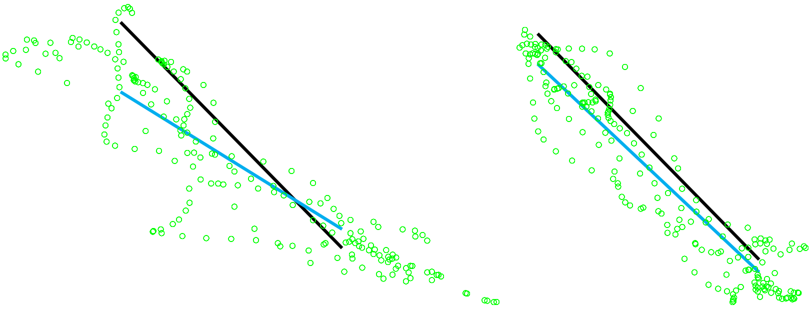
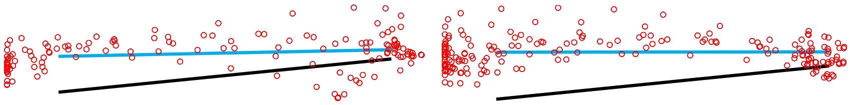
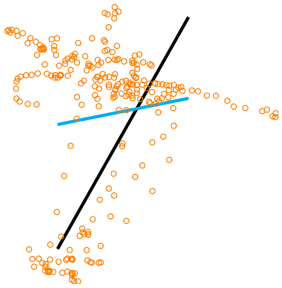

(e) Green can - no strategy (f) Green can - strategy

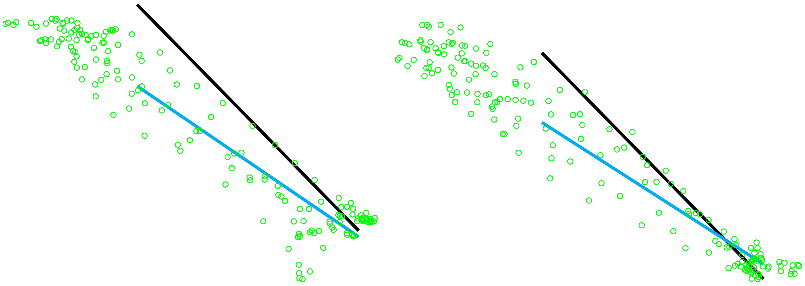
Figure 6.20: Patient A, approximate (light blue) and optimal (black) trajectories, and hand positions, based on the can color (dots).



(a) Red can - no strategy (b) Red can - strategy

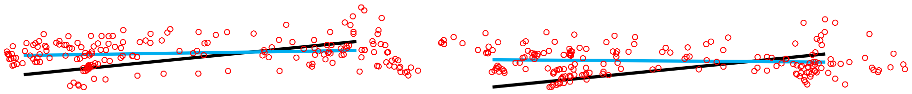


(c) Orange can - no strategy (d) Orange can - strategy

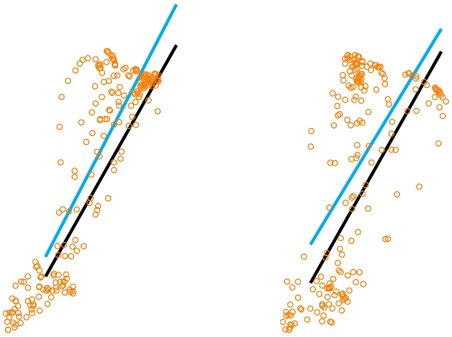


(e) Green can - no strategy (f) Green can - strategy

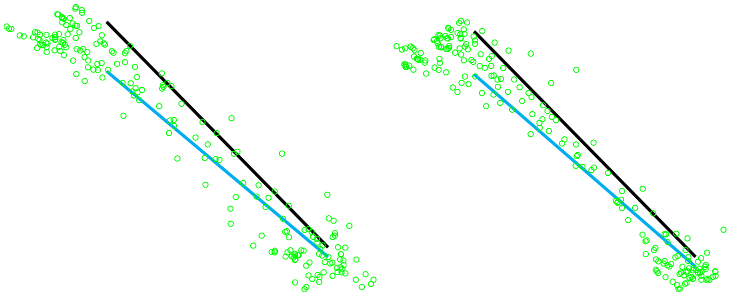
Figure 6.21: Patient B, approximate (light blue) and optimal (black) trajectories, and hand positions, based on the can color (dots).



(a) Red can - no strategy (b) Red can - strategy

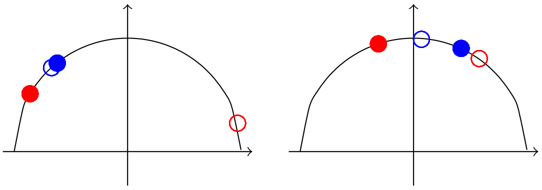


(c) Orange can - no strategy (d) Orange can - strategy

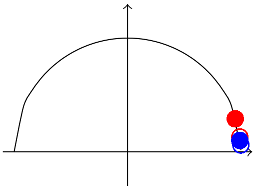


(e) Green can - no strategy (f) Green can - strategy

Figure 6.22: Subject 1, approximate (light blue) and optimal (black) trajectories, and hand positions, based on the can color (dots).



(a) Shoulder ROM (b) Elbow ROM



(c) Trunk ROM

Figure 6.23: Patients A and B shoulder, elbow, and trunk ROMs visualizations. Red marks refer to Patient A, blue marks refer to Patient B. Empty circles are for sessions without adopting the weighting strategy, full circles are for sessions where the weighting strategy was adopted.

6.5 Application 4: Frailty

The application of ReMoVES in the context of frailty[4] is disclosed with regards to the exergame Equilibrium Paint.

Some peculiar features are defined and extracted from the patients' game sessions in order to provide a picture of the general conditions of the considered population. The definition of the considered features was inspired by works in the literature such as [192, 193]. Similarly to the Section 6.4, they were computed from the spatial coordinates of joints in Figure 4.2.

The main indicator is the number of sit-up occurrences (NSU) during the 30 s duration of the test. This is computed by analyzing the trajectory of the spine middle joint (joint 7 in Figure 4.2) along the vertical axis. Each peak of such a trajectory represents a sit-up. For the population under analysis, the average NSU is $N\bar{S}U = 4.5$.

Peak detection also allows for separating the ascending and descending phases during activities. They are identified as the parts of the trajectory between a local minimum of the spine middle height and the following peak, and between a peak and the following local minimum of the spine middle height, respectively. In this fashion, it is possible to in-depth analyze both phases by computing ad hoc indicators.

The first feature that is introduced is the upper-body flexion angle (UBFA), which represents the angle of flexion of the trunk and is computed as

$$UBFA = \arctan \frac{z_2 - z_7}{y_2 - y_7}. \tag{6.9}$$

The UBFA is maximal when the player is in a standing position, and reaches values of approximately 90° when sitting. In addition, other values are present that represent the intermediate phase from a sitting to a standing position and vice versa. For standing up,

---

[4] The studies on frailty were conducted in collaboration with Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, and Maternal and Children's Sciences (DINOGMI) - Università degli Studi di Genova, Clinica di neuroriabilitazione - IRCCS Ospedale San Martino, Centro di riabilitazione srl and Struttura Complessa Recupero e Rieducazione Funzionale - ASL3. In addition to the paper [138], some results were also presented to the medical congresses XVIII, XIX, and XXI Congresso Nazionale della Società Italiana per lo Studio dello Stroke.

the player should move forward, which results in a decrease in sitting UBFA.

Similarly to the UBFA, the indicator of the lower-limb flexion angle (LLFA) represents the knee angle, and can be computed for both the left and the right limb. It is defined as

$$LLFA = 180 + \theta_{femur} - \theta_{tibia}. \tag{6.10}$$

where $\theta_{femur} = \arctan \frac{z_{20} - z_{19}}{y_{20} - y_{19}}$, $\theta_{tibia} = \arctan \frac{z_{22} - z_{20}}{y_{22} - y_{20}}$ for the left limb and $\theta_{femur} = \arctan \frac{z_{21} - z_{17}}{y_{21} - y_{17}}$, $\theta_{tibia} = \arctan \frac{z_{24} - z_{21}}{y_{24} - y_{21}}$ for the right limb. Variation in this angle for both the left and the right limb is similar to the trajectory of the spine middle joint (see Section **??**).

During this activity, patients may adopt erroneous behavior such as moving the shoulders or hips. Hence, it is important that therapists supervising the rehabilitation are informed about these compensatory movements. Regarding shoulder movement, the upper-body twist angle (UBTA) depicts the angle of the line joining the shoulders (joints 3–5 in Figure 4.2) on the axial plane:

$$UBTA = \arctan \frac{y_5 - y_3}{x_5 - x_3}. \tag{6.11}$$

Hip displacement is calculated on the basis of the anteroposterior and mediolateral displacement of the Center of Mass (COM). The COM is defined as the middle point between the right and left hips (joints 17 and 19 in Figure 4.2, respectively) and spine middle (joint 7 in Figure 4.2):

$$COM = (\bar{x}, \bar{y}, \bar{z}) = \left(\frac{x_{17} + x_{19} + x_7}{3}, \frac{y_{17} + y_{19} + y_7}{3}, \frac{z_{17} + z_{19} + z_7}{3} \right). \tag{6.12}$$

Hence, indicators COM anteroposterior (A-P) movement and COM mediolateral (M-L) movement depict COM positions on the axial plane.

To conclude, upper-frame velocity (UfV) is the velocity of motion in either the ascending or descending phase. For one ascending phase, it is computed as

$$UfV_{up} = \frac{z_{peak} - z_{localmin}}{time_{peak} - time_{localmin}}. \tag{6.13}$$

Similarly, in the descending phase, it is

$$UfV_{down} = -\frac{z_{localmin} - z_{peak}}{time_{localmin} - time_{peak}}. \tag{6.14}$$

More generally, all aforementioned features are separately computed in the ascending and descending phases in order to provide a fragmented and specific analysis of the patients' sessions.

A real-world case study involved 13 frail elderly people (6 females and 7 males) with an average age of 82.3 ± 6.2 who participated several times to the rehabilitation sessions via ReMoVES . This was a preliminary feasibility study to evaluate the possible use of ReMoVES in a real-world scenario. Feedback from the present work drives further applications involving more patients. Participants reported that they felt safe while playing the game, and there were no adverse events while playing. Most of the patients stated that they enjoyed this extra activity, asking the clinical staff to participate more frequently. An interesting social interaction developed among the participants, who enjoyed watching others carry out the activities.

The implicit analysis of the activity performed by the involved patients is presented. Mean values of the proposed indicators were collected, and their coherence with already published results was statistically tested. In addition, aimed at enabling deeper analysis of each game session, a graphic visualization of the indicators along the time dimension is shown. Such graphs are provided to therapists via the application layer, so that clinical staff analyze both summary statistical indicators and patient performance during the whole session. In this fashion, even some erroneous movements or loss of energy, which may be limited to a short period of time, can be noted by the medical specialists, leading to a complete and deep clinical picture of the patients.

The average features of the available population are summarized in Table 6.9. Negative values for the UBTA indicate that the left shoulder was put forward while practicing the activity.

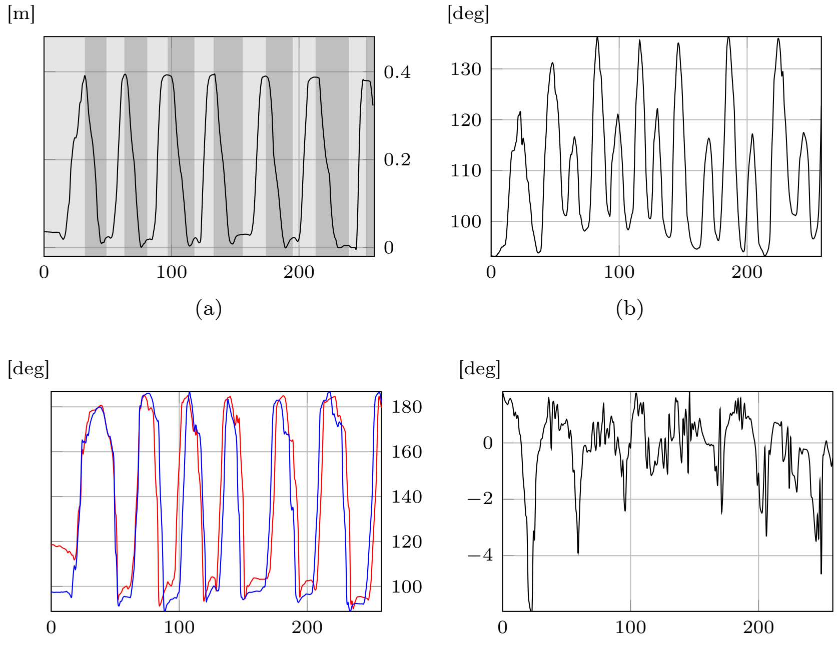
Table 6.9: Mean feature values. NSU, number of sit-up occurrences; UBFA, upper-body flexion angle; LLFA, lower-limb flexion angle; COM, center of mass; A-P, anteroposterior; M-L, mediolateral.

Feature	**Mean Value**
NSU	4.5 ± 1.5
Stand UBFA range (deg)	79.92 ± 6.71
Sit UBFA range (deg)	79.35 ± 8.15
Stand LLFA (deg)	131.16 ± 17.28
Sit LLFA (deg)	134.31 ± 16.94
Stand UBTA (deg)	-0.67 ± 1.91
Sit UBTA (deg)	-0.59 ± 1.91
COM stand A-P (cm)	0.36 ± 0.09
COM sit A-P (cm)	0.52 ± 0.61
COM stand M-L (cm)	0.08 ± 0.02
COM sit M-L (cm)	0.07 ± 0.03
UfV_{up} (m/s)	0.12 ± 0.06
UfV_{down} (m/s)	0.07 ± 0.02

To address the coherence of the derived data with respect to the literature, the results of [192, 193] were considered for the discussion. In [192], the indicators standing and sitting COM A-P, standing and sitting COM M-L, UfV_{up}, and UfV_{down} were calculated with respect to a population of healthy elderly individuals (mean values were 0.01, 0.03, 0.03, and 0.04 cm, and 0.78 and 0.71 m/s, respectively). Hence, a statistical test was performed to verify the assumption that the indicator values in [192] depicted a better general health condition than the ones deduced for the population under analysis. A one-tailed t-test was used, and the assumption was confirmed with p value < 0.01.

In addition, the authors in [193] showed mean values for the range of UBFA in both the ascending and descending phases in a population of frail elderly persons. Via a two-tailed t-test, the assumption that the mean ranges of UBFA in [193] and in the present work were equal was verified with p value < 0.01.

Therapists were also provided with graphs depicting all game sessions, delivering more comprehensive informative content than that in the mean or range indicators. An example of the graphic representation available on the therapist client is shown in Figure 6.24. In particular, Figure 6.24a depicts the trajectory of the COM and peaks; hence, corresponding standing positions were visible. In particular, parts with a light-gray background are for the ascending phase, and parts with a dark-gray background represent the descending phase. Figure 6.24b shows UBFA values during the session. Figure 6.24c shows LLFA values for both the right and the left limb during the session. The trend of this chart is very similar to that of the COM. A standing position also requires limbs to be fully extended, corresponding to the peaks of the LLFA indicators. Figure 6.24d presents the shoulder twist on the axial plane during the session. Lastly, COM A-P and COM M-L displacements are depicted in Figure 6.25 on the transverse plane. Reduced lateral displacement in the second graph with respect to the first suggests that the patient stabilized themselves while playing.



(a) (b)

(c) (d)

Figure 6.24: Graphs from Equilibrium Paint exergame. These graphic representations are available for clinical staff, so that deeper analysis is enabled throughout the whole session. (**a**) COM; (**b**) UBFA; (**c**) LLFA; (**d**) UBTA.

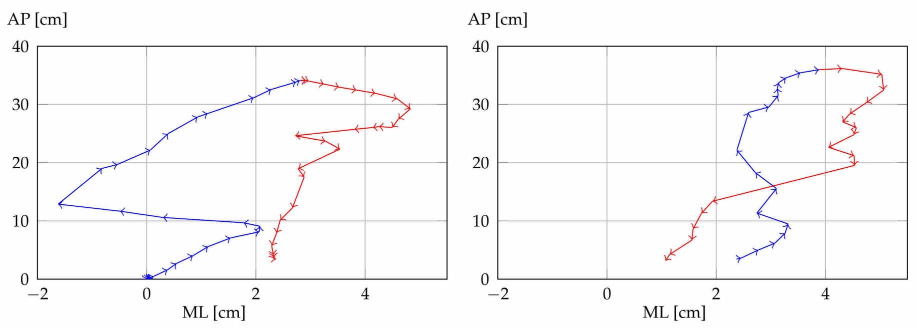


Figure 6.25: Representations of A-P and M-L movements of COM in two consecutive ascending and descending phases. Blue lines, ascending phases; red lines, descending phases.

The COM trajectory in Figure 6.24a shows that the patient performed a smooth movement with no particular pauses. The resulting regular path means that the patient did not experience particular fatigue and managed to control their motion. So, by only considering such a graph, a therapist would say that the patient's performance was fairly good. However, Figures 6.24 b,c, for UBFA and LLFA, respectively, depict incomplete movement. Indeed, the patient is supposed to reach maximal extension while standing, namely, the maximal values of UBFA and LLFA (corresponding to COM peaks) should reach approximately 180°. While LLFA satisfies such a requirement, meaning correct leg extension, the maximal values of UBFA were around 130°, denoting that the patient remained bent forward when standing. Figure 6.24d for UBTA depicts that shoulder rotations were very small, denoting correct movement (the patient is required to preserve shoulders in the frontal plane, i.e., without trunk rotations). To conclude, graphs in Figure 6.25 depicting A-P and M-L movements show that the patient was not laterally significantly displaced (about 2 cm), confirming the correct execution of the exercise apart from the vertical trunk extension.

This shows how multidimensional data can provide the clinical staff with precise information. This is very important for reliable remote monitoring, by which small or partially erroneous behaviors can also be detected and corrected.

6.6 Application 5: Balance assessment

The exergames involved in the present study are Equilibrium Paint, Push Box, and Hot Air, i.e., the ones for training standing, mediolateral, and antero-posterior balance.

For each game the appropriate joints are traced in 3D to observe the movement patterns. In this way, a number of discrete signals are obtained, from which it is possible to extract in a synchronized way with respect to the execution of the exercise the most significant features regarding the amplitude of the movement, the angles, etc.

Once the signals are acquired, a series of operations are applied to process them and extract the information provided. The type of noise that primarily affects the signal is impulsive. So a nonlinear filtering operation is required. However, the classic median filter does not work properly in the current situation, due to the large peak density. As a result, adaptive filtering, consisting of a combination of outlier detection and median filtering, is designed to guide the de-noising operation based on the statistical properties of the signal. In particular, once the outliers are detected, they are replaced by the median value in a window of dimension five, in order to remove the noise while preserving the original signal as much as possible.

Then, to proceed with a reliable analysis, a signal segmentation step is necessary, aimed at focusing the subsequent extraction of the characteristics on stationary signal portions. The result is the decomposition of the complex movement into primitive ones. Therefore, the segmentation process is specific to each exergame, as it is based on the required movement and the related joints of interest.

Filtering

The initial and final fifteen samples have been removed as they correspond to the transient part of the signals. Second, impulsive behaviors were observed in the three dimensions of signals, possibly due to the temporary loss of Kinect joint tracking. Since large or high-frequency calibration errors can significantly alter the movement of signals and lead to errors in the measurement of clinical indicators, filtering techniques had to be applied prior to their calculation. The way to deal with outliers was to remove values whose distance from the mean of the signal was greater than three times its standard deviation. This condition arose from the idea of thinking of impulsive noise as a sample that exceeds the 0.9 percentile in the signal's probability density function.

Since the signal statistical model is not perfectly known, the more conservative Chebyshev inequality approach [194] is applied to obtain the threshold value:

$$P(|x - \mu| \leq \lambda\sigma) \geq 1 - 1/\lambda^2 \tag{6.15}$$

with $1 - 1/\lambda^2 = 0.9$, x being the random variable associated to a generic signal sample, described by μ and σ^2, which are respectively its mean and variance. In such a case, outlier values are the ones differing more than three times the standard deviation from the mean value.

An outlier value detected is replaced with the corresponding value by the median filter applied to the non-abnormal samples.

An example of filtering operation is provided in the following (Figure 6.29) with regards to a member of the impaired population.

Motion segmentation

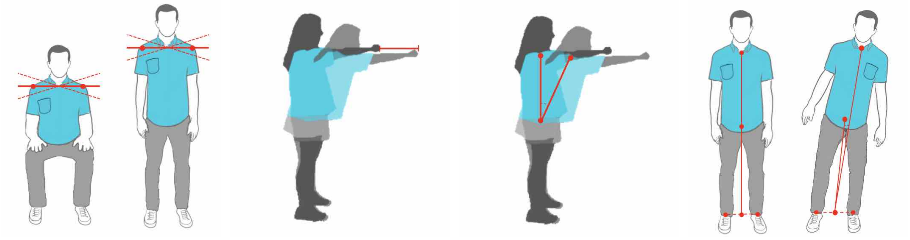
The extracted features have been obtained by considering blocks of data of the sessions, which will be hereinafter referred to as *cycles* or *segments*. Such a motion segmentation is performed to enable a deep analysis of the activity, and it depends on the exergame, namely it is driven by the movement pattern required by the game task.

Concerning Equilibrium Paint, the local maxima and minima of the body center of mass are identified to discern ascending and descending phases. Indeed, features are computed separately with respect to ascending and descending phases, in order to highlight also in which of these cases the patient may have issues.

As for Push Box, cycles are defined based on the forward elongation. The beginning of the elongation phase is identified when the ordinate of the fingers exceeds the threshold ordinate, calculated as the midpoint between spine base and spine mid joints. Then, the maximum point reached by the fingers along the frontal axis is set to identify its end.

Finally, in Hot Air, the cycles are defined on the basis of the center of gravity corresponding to the initial standing position. The signal portions correspond to movements to the right or left with respect to the initial situation. Numerically, the neutral reference value is the barycentric mean on the lateral axis, thus being an adaptive threshold value obtained on the basis of each session.

In Figure 6.26 the cycles for each exergame are depicted.



(a) Cycle in EP and Features 1, 2.
(b) Cycle in PB and Feature 3.
(c) Features 4, 5, from PB and HA respectively.
(d) Cycle in HA and Feature 6.

Figure 6.26: Representation of cycles and features in each exergame.

Also, Figure 6.27 shows an example of the result of the signal segmentation, with regards to the Hot Air exergame. Light gray areas are for the left displacement, while dark gray ones are for right displacement.

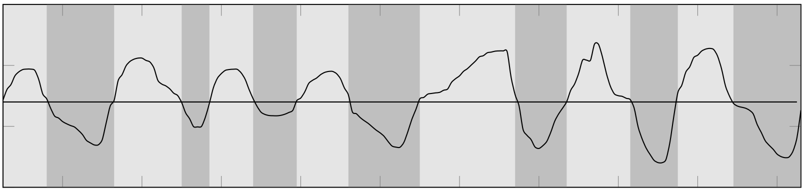


Figure 6.27: Example of movement segmentation in Hot Air exergame.

It has been shown that the features considered here significantly describe the control population, since they become homogeneous when referred to a homogeneous population. The results demonstrate good repeatability, reproducibility, and reliability in observing the characteristics of interest. The list of such features follows.

- *Feature 1* and *Feature 2* - Equilibrium Paint (ascending and descending phase respectively): the average angle between the segment joining the shoulders and the horizontal line in the coronal plane. Figure 6.26a depicts such features.
- *Feature 3* - Push Box: the elongation, i.e., the difference between the maximum and the minimum of the fingers coordinate in the anteroposterior axis. In Figure 6.26b a representation of such feature is provided.
- *Feature 4* - Push Box: the range of the anteroposterior inclination of the trunk. Figure 6.26c shows the presented feature.
- *Feature 5* - Hot Air: the range of the anteroposterior inclination of the trunk. The present feature is the same as Feature 4, but in a different exergame. It is shown in Figure 6.26c.
- *Feature 6* - Hot Air: the difference between the maximum and the minimum angles between ankle, the middle point between the ankles, and the top of the spine. Figure 6.26d represents Feature 6.

An example of feature visualization in time is provided in Figure 6.30, with respect to an impaired subject.

Three types of subjects have been involved in the present study. A control group of twenty-four healthy subjects was considered. They are clinical specialists from La Colletta Hospital (ASL3 Sistema Sanitario Regione Liguria, Italy). Their sessions are used to define the benchmark values for further comparisons, based on the most significant indicators. Then, a group of twenty-five patients was included in the study. They are frail elders presenting several clinical pictures, with different pathologies and different levels of disease. In general, most of them are post-stroke patients, showing hemiparesis as a consequence. In addition, some of them presented fractures and orthopedic diseases affecting the lower limbs. The game sessions by each of them is compared with the control group and significant differences are found. However, their conditions were too spread to define a single group of impaired subjects and define the benchmark values for patients. Such a goal will be accomplished with further studies. Finally, a patient affected by a neuro-degenerative disease was taken into account. He practiced the activities both in the hospital and remotely at home, thus a set of sessions was collected, enabling some longitudinal reasoning about his treatment and suggesting a possible usage of the ReMoVES system.

Each participant had some preliminary training with the system, aimed at familiarizing with it. Then, the considered sessions were performed and collected.

The homogeneity of the available population was accounted via an ANOVA test. The test was performed separately with respect to each feature previously presented.

In fact, the first ANOVA tests have been performed on the control group, aimed at proving the homogeneity of such a sample with regards to the involved features. In this fashion, the population of healthy subject can be considered a *good sample.*

Then, the impaired population was considered. Here, the ANOVA tests were used to prove that the homogeneity of the control group fails whenever a patient is added to the healthy population. This proves the significant difference between patients and control group.

Finally, the tests were performed taking into account the longitudinal sessions by a patients, following a remote rehabilitation process, aimed at assessing his activity based on the control group sessions.

In all the case, the significance levels of the ANOVA tests were set $\alpha = 0.05$.

With regards to the control group, the ANOVA tests yielded to accept the null hypothesis for all the features $1 - 6$, proving the homogeneity of the considered population. The average values of such indicators are summarized in Table 6.10.

Given that the control group have proven homogeneity, such a population can be considered a benchmark denoting reliability, robustness, and reproducibility of the ReMoVES system.

A visualization of a feature involved is depicted in Figure 6.28. Here, the values of Feature 6 for each subject are shown, with the red dots representing the mean value for the person and the horizontal

Table 6.10: Average values of the considered features for the control group.

Feature	**Average**
Feature 1 - EP	1 deg
Feature 2 - EP	2 deg
Feature 3 - PB	34 cm
Feature 4 - PB	63 deg
Feature 5 - HA	3 deg
Feature 6 - HA	36 deg

line showing the mean value of the whole population.

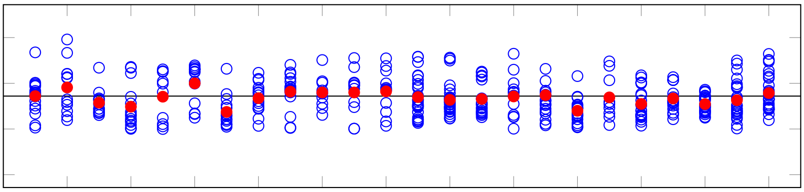


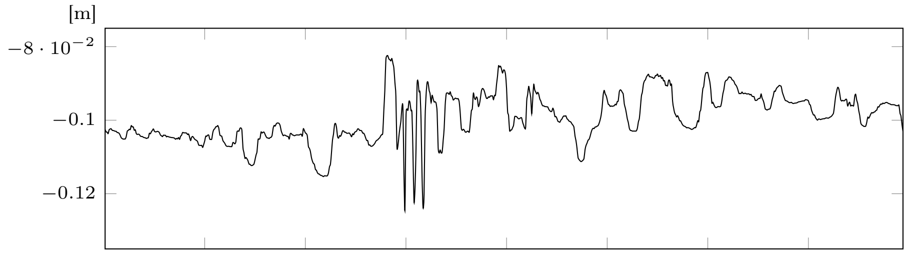
Figure 6.28: Scatter plot Feature 6 - control group.

Patients involved presented different conditions, and indeed the homogeneity of the considered population was not proven. As this is not surprising, the goal of the present experiment was to show that each patients session is significantly different from the benchmark control group, based on the considered features.

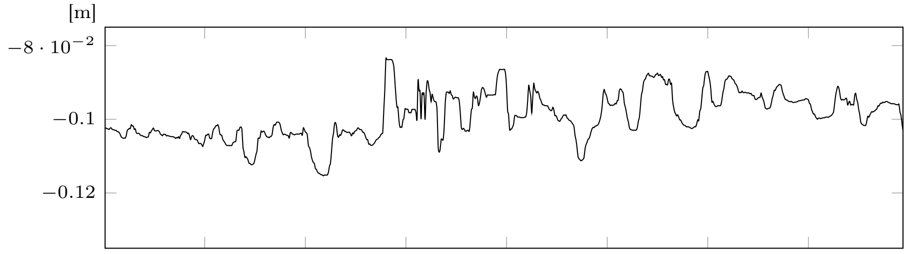
Each patient was added to the healthy population (one at time) and the ANOVA test yielded to reject the null hypothesis of homogeneity. This proves that the homogeneity of the control group fails as long as a patient is added to the population, meaning that patients are significantly different from the healthy subject.

The most important aspect of such results is that, based on the benchmark population, it is possible to monitor the activity of patients, evaluating whether a person is correctly performing - as a healthy subject - or not. Such a possible application is disclosed in the following experiment with respect to a patient affected by a neuro-degenerative disease following a home-based treatment.

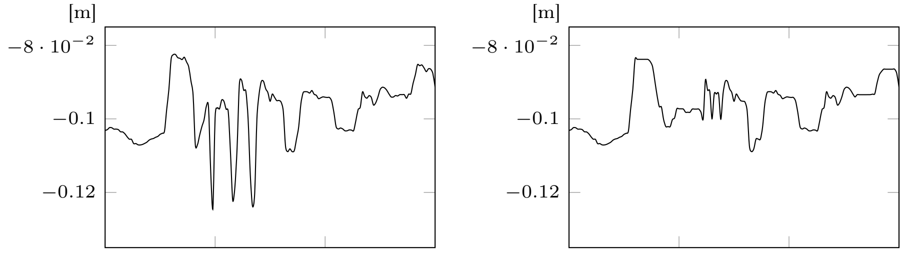
In addition, the effect of the filtering operation are here shown. Taking into account an example of signal traced by the Kinect sensor, Figure 6.29 depicts how signal quality is enhanced by the proposed filtering approach, with also a focus on a detailed frame of the signal.



(a) Raw signal



(b) Filtered signal



(c) Raw signal detail

(d) Filtered signal detail

Figure 6.29: Signal filtering action, HA exergame - full signal and frame detail.

Finally, a visualization, of a derived signal, i.e., a feature, is provided in Figure 6.30.

As a second application of the defined benchmark data, a patient affected by a neuro-degenerative disease was considered. He had been following a home-based rehabilitation treatment which included Equilibrium Paint and Hot Air exergames, and lasted twenty-

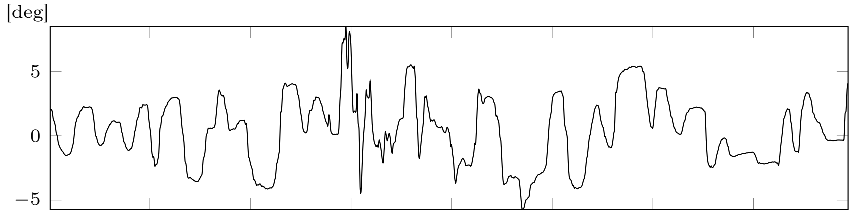


Figure 6.30: Feature 6 visualization for a patient.

four days. Push Box activity was not practiced as the clinical staff believed it was not safe, based on patients conditions.

Similarly to the procedure followed for assessing the impaired population, each of the twenty-four sessions performed by the patient was added to the healthy population one at time. In this fashion, the clinical staff could evaluate whether the patient was either well- or poorly performing, even though they were not directly assisting the rehabilitation session.

Similarly to the procedure followed for assessing the impaired population, each of the twenty-four sessions performed by the patient was added to the healthy population one at time. In this fashion, the clinical staff could evaluate whether the patient was either well- or poorly performing, even though they were not directly assisting the rehabilitation session.

The ANOVA test yielded to accept the hypothesis of homogeneity when referring to the movement required by the exergame Hot Air. Conversely, the activity for Equilibrium Paint exergame resulted impaired.

A discussion on the outcomes of such an analysis is provided in the following.

The characterization of the healthy population allows for extracting some information related to how they perform the activity. In particular, the fact that the homogeneity of the sample was granted for the features previously described leads to focus on such movements also for the analysis of sessions by patients. Indeed, the homogeneity of the healthy population was a prior assumption as they all have similar characteristics such as age, strength, height, weight, and also the share the same professional occupation. As a result, one can infer that the features supporting the homogeneity of the population are the ones that are mostly significant for describing the activity.

Starting from Equilibrium Paint, the features granting homogeneity is the one describing the shoulder displacement in the coronal plane. In particular, also taking into account the result in Table 6.10, the line joining shoulder can be considered horizontal in both the ascending and descending phases. This denotes a good control of the standing and sitting movements, which are performed without any compensation. It worth noting that the numbers of sit-ups is not a significant feature for the control group. This may also be due to the small sample size considered, but, in general, one can conclude that the number of sit-ups is too variable to be a very good index. Instead, the posture when practicing such an activity results in being the very representative feature for the sit-to-stand activity.

Concerning the exergame Push Box, it emerged that the most significant features are elongation and anteroposterior trunk inclination. Of course, as long as the movement is correctly performed, such indices are strictly linked, since large elongation can be reached by bending the trunk. In general, the healthy subjects share a similar extension ability, which is granted by similar balance ability and equilibrium. Furthermore, control subjects are not much afraid of falling, as their conditions are likely to enable them to recover from a loss of balance. Also, it is worth noting that the aim of the game is to make the maximum possible stretch, keeping the lower limbs still by tilting only the trunk. For this reason, the inclination feature is particularly important, so that in general, the larger the angle, the more correct the movement. In fact, smaller angles could mean that the movement was accomplished by tilting the lower limbs instead of the trunk, which significantly increases the risk of falling.

Finally, in Hot Air exergame, two are the most significant features. The former one is the anteroposterior trunk inclination, as in the Push Box exergame. It is noteworthy that, even though such a property is strictly linked to the balance, it may seem surprising in this context as the present exergame requires a lateral displacement. However, healthy subjects are all able to avoid anteroposterior motions (Feature 5 in Table 6.10), and this results in being a discriminatory ability.

The latter is a more expectable feature, as it describes the lateral displacement of the patients which is the required movement in Hot Air exergame. The reader may notice that an easy compensation which allows for keeping playing it to displace hips, without training the mediolateral balance. However, since the present features take into account an angle referred to ankle, hips, and top of the spine joints, it considers the correct and un-compensated movement.

All of these features highlighting good balance ability for the control group are likely to fail when considering patients.

Indeed, in Equilibrium Paint exergame, patients showed significant differences in shoulders alignment in both the ascending and descending phases. It is mainly due to two reasons. The former is that patients considered presented asymmetric impairments due to either post stroke outcomes or orthopedic impairments on the lower limbs. As a result, both standing and sitting movements were impaired and asymmetric, as they weighted more on the dominant limb and hemi-body, yielding to a shoulders displacement in the coronal plane. The latter is due to a common strategy used by hemiparetic patients for fulfilling the sit-to-stand task. It consists on placing the unaffected foot slightly further ahead of the other, aimed at pushing more on that. On the one hand, this positioning allows a better standing motion; on the other hands, it enhances the asymmetry of the movements, affecting also the shoulders alignment which, in summary, was significantly different from the one in the control group.

Then, in Push Box exergame, the considered features address the ability of forward bending. Patients present poor balance ability due to their conditions, which affect the possibility to elongate without the risk of falling or, at least, fearing to fall. As a results, all the patients were keen on performing a small elongation, aimed at keeping themselves in a sort of *balance comfort zone*, granting them to avoid the risk of falling and its perception. Therefore, their movement results very limited and significantly different with respect to the one by a healthy subject.

To conclude, Hot Air exergame yielded to two other considerations related to balance ability in patients. As for Feature 5 measuring frontal displacement, it emerges that patients are likely to perform also a significant anteroposterior movement even though it is not required by the exergame. Actually, it denotes an unsteadiness in the standing posture, which yields them to swing thus acting significantly different from the control group. Then, Feature 6 showed a limited mediolateral movement. In this case, similarly to the forward elongation in the Push Box exergame, patients preferred avoiding large movements which may stress their balance and make them fall.

The final part of the discussion is for the study on the home-based rehabilitation sessions.

Starting from the Hot Air activity, the patient proved good balance abilities which are actually in line with the one by the control group. In particular, Figure 6.31 depicts a short summary of the sessions, based on the defined features. Here, full circles represent Feature 5, while empty circles are for the Feature 6. Whenever the ANOVA test proved homogeneity in the joint population of healthy subjects plus the current patient's session, a blue circle was included in the graph, placed over a discriminator line. Conversely, if the result of the ANOVA test is the rejection of the homogeneity hypothesis, a red circle is place under the discriminator line.

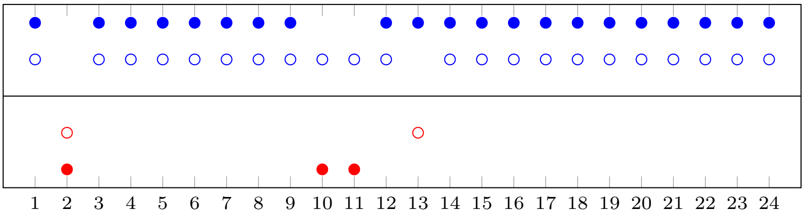


Figure 6.31: Performance of the patient during the longitudinal assessment. Blue signs are for sessions where the features are significantly similar to the ones by the control group. Conversely, red signs are for sessions when features significantly differ from the control group. Full circles are for the Feature 5, empty circles are for the Feature 6. In general the patient well-performed the Hot Air exergame, with respect to the control group.

However, when taking into account Equilibrium Paint, it can be noticed that patient's sessions are significantly different from the control group. The rationale behind Figure 6.32 is the same as in Figure 6.31.

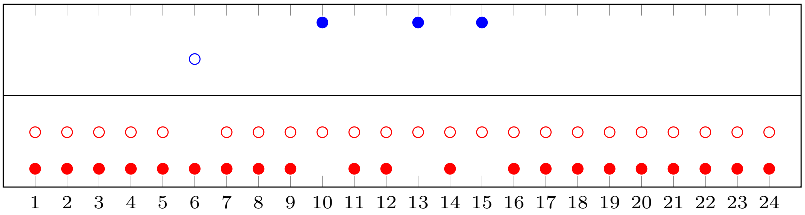


Figure 6.32: Performance of the patient during the longitudinal assessment. Blue signs are for sessions where the features are significantly similar to the ones by the control group. Conversely, red signs are for sessions when features significantly differ from the control group. Empty circles are for the Feature 1, full circles are for the Feature 2. In general the patient poorly performed the Equilibrium Paint exergame, with respect to the control group.

In this case, it could be interesting for the clinical staff to verify whether the patient is actually worsening his activity, or if the patients is at least maintaining his ability, even if he is not improving and getting similar to healthy subjects in performance. Indeed, it is actually rare that neuro-degenerative patients significantly improve their activity, while maintenance is often a more concrete and reachable goal of the rehabilitation. Conversely, a general motion pattern is likely to be kept and can be monitored by the clinical staff.

To this purpose, it is possible to consult the values of the involved indicator, and take into account their variation in time, i.e., longitudinally. Figures 6.33 and 6.34 depict such feature values variation in time.

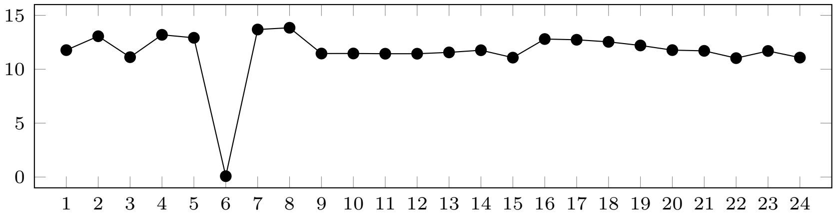


Figure 6.33: Feature 1 variation along session

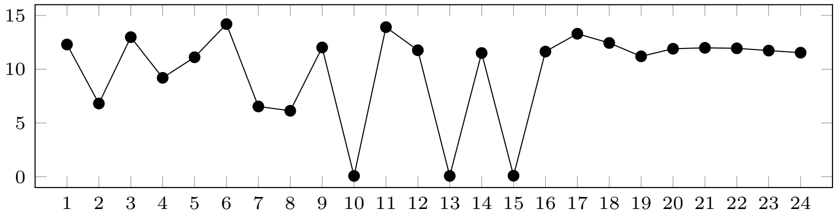


Figure 6.34: Feature 2 variation along session

From such graphs, it can be noted that the general trend of the features is constant, apart from the Sessions where the patients performed as the control group. However, since such sessions are a few, they can be considered as occasional outstanding performance rather then a sign of improvement.

Even though in this case the longitudinal study was conducted with regards to the control group, an approach customized on the patient can be adopted. Indeed, patients' sessions could be characterized while in the facility, to monitor home-based activity using themselves as reference values.

CHAPTER 7

Conclusion and future perspectives

The goal of the present thesis was to propose novel approaches for biomedical image and signal processing, aimed at fostering the development of digital and enabling technologies in the medical domain.

Two major topics have been covered, i.e., diagnostic imaging and tele-rehabilitation. Indeed, such application have been acquiring relevance for the last years, and also the current pandemic emergency is stressing the need for agile and efficient medical technologies.

Diagnostic imaging and tele-rehabilitation are linked procedure from the points of view of both the medical practice and the technologies aspect. As for the former, it is clear how rehabilitation is crucial with regards to many pathological conditions, which may be diagnosed based on image findings. As the latter, the development of some technology for diagnostic purposes is intended to deliver information useful to overcome the pathology-related issues, where (tele-)rehabilitation can act. In addition, the development of web-based technologies enables an easy access to both diagnostic and disease-management related information, and are of great interest nowadays.

A possible naive pipeline starting from diagnostic imaging and leading to recovery, passing through technology-enabled rehabilitation is shown in Figure 7.1.

The deployed techniques and conducted studies were exploited by referring to the two research themes in parallel, promoting the study of the full medical process for a single pathology as a future development.

For what concerns diagnostic imaging (Chapters 3 and 5), the major contributions of the present work are referred to the context of image segmentation. In particular a novel technique for unsupervised image segmentation was presented. It is an ensemble of graph- and statistical signal processing, aimed at the segmentation of regions of interest in the scene, which are meaningful for some application, e.g., the detection of carpal bones in MR images of the

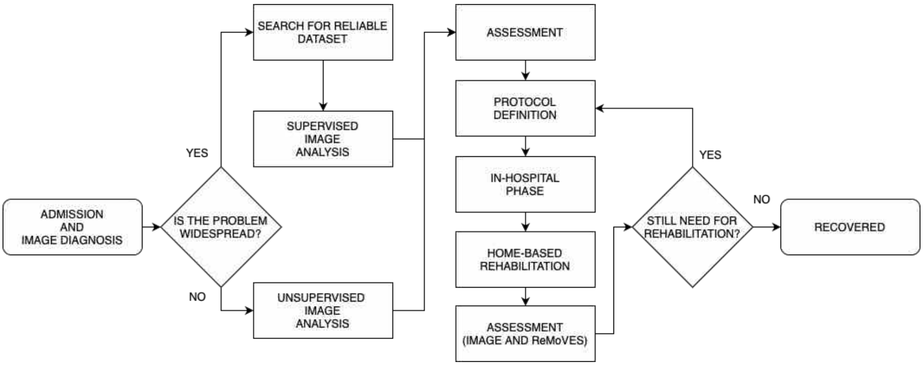


Figure 7.1: General pipeline for the medical practice based on the proposed techniques.

wrist.

In addition, the problem of human bias in automatic medical image analysis is tackled. In such a context, some operations are suggested, which may be useful either for standardizing the reference data or for making the evaluation process objective and independent on the raters. The applications taken into account are the unsupervised evaluation of segmentation results, and the introduction of a pre-processing phase prior manual annotations, aimed at reducing raters dependence of ground-truth images. Also, this last study suggests a glimpse into the development some web-based tool for medical images annotation, aimed at enlarging the pool of annotators thus making the labeled data more reliable.

With regards to tele-rehabilitation (Chapters 4 and 6),, ReMoVES was involved in studies for several pathologies, including Systemic Sclerosis, Unilateral spatial neglect, post stroke, and frailty. Data from patients experiencing such diseases were analyzed with SoA methods, for classification purposes and for defining proper indicators to summarize game sessions based on the considered pathology.

While the paragraphs above are meant to recapitulate the results of the present work, it is interesting to spend some words on the research framework which the present studies are contextualized in, and on the aspect that they may affect.

The research activity is compliant with the National Strategy of Intelligent Specialization (*Strategia Nazionale di Specializzazione In-*

telligente - SNSI) and the National Research Program (*Programma Nazionale per la Ricerca* - PNR), and coherent with Law 240/2010, with the aim of favoring innovation and interchange between the world of research and the world of production and the qualification of the contribution of the activities research in the fields of innovation (Law 240/2010, art.24, co. 3 and subsequent amendments).

Also, the current Covid-19 pandemic has highlighted the importance of applications such as telemedicine, remote patient monitoring, and interactive medicine so that they are acquiring a crucial position in the novel medical practice.

To conclude, also from an industrial point of view, the healthcare technology sector is significantly expanding, as innovative medical technologies are able to deliver substantial new opportunities to those healthcare facilities implementing them. Indeed, in a world where disruptive digital technologies are revolutionizing almost everything, also patient care could benefit from such an innovation process, yielding to better working condition for the clinical staff, and thus, to an improved patient management and experience.

Appendix A

PhD activities

A.1 Publications

- Trombini, M., Morando, M., Berti Riboli, E., Zanna, M., Rudelli, A., Dellepiane, S. “The SPeRA web Portal for Cooperation Associations and Projects in Africa.” IEEE Transactions on Technology and Society (submitted).

- Trombini, M., Solarna, D., Moser, G., Dellepiane, S. “A goal-driven unsupervised image segmentation method combining graph-based signal processing and Markov random fields.” Pattern Recognition - Elsevier (submitted).
- Trombini, M., Dellepiane, S. “Medical Images Segmentation: how raters’ experience may affect the quality of reference information.” Computational Intelligence and Image Processing in Medical Applications. World Scientific Publishing (to appear).
- Trombini, M., Ferraro, F., Iaconi, G., Vestito, L., Bandini, F., Mori, M., Trompetto, C., Dellepiane, S. “A Study Protocol for Occupational Rehabilitation in Multiple Sclerosis.” Sensors (2021); 21(24), 8436. https://doi.org/10.3390/s21248436
- Trombini, M., Ferraro, N., Nardelli, A., Vestito, L., Schenone, G., Mori, L., Trompetto, C., Dellepiane, S. “On the performance assessment during the practice of an exergame for cerebellar ataxia patients.” 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 5747-5751. https://doi.org/10.1109/EMBC46164.2021.963028
- Trombini, M., Ferraro, F., Dellepiane, S. “On the use of boundary gradient for the analysis of MR wrist bones volumes segmentation.” Proceedings of the 2021 International Conference

on Intelligent Medicine and Image Processing, 2021.
https://doi.org/10.1145/3468945.3468950

- Ferraro, F., Trombini, M.*, Morando, M., Doveri, M., Bianchi, G., Dellepiane, S. "Exergame for systemic sclerosis rehabilitation: a pilot study." Advances in Computer Vision and Computational Biology. Transactions on Computational Science and Computational Intelligence. Springer, Cham.
https://doi.org/10.1007/978-3-030-71051-4_22

- Trombini, M., Ferraro, F., Manfredi, E., Petrillo, G., Dellepiane, S. "Camera Color Correction for Cultural Heritage Preservation Based on Clustered Data." Journal of Imaging (2021) 7(7):115.
https://doi.org/10.3390/jimaging7070115

- Ferraro, F., Trombini, M., Truffelli, R., Simonini, M., Dellepiane, S. "On the Assessment of Unilateral Spatial Neglect via Digital Tests." 10th International IEEE EMBS Conference on Neural Engineering.
https://doi.org/10.1109/NER49283.2021.9441471

- Trombini, M., Ferraro, F., Morando, M., Regesta, G., Dellepiane, S. "A Solution for the Remote Care of Frail Elderly Individuals via Exergames." Sensors (2021); 21(8):2719.
https://doi.org/10.3390/s21082719.

- Borro, P., Ziola, S., Pasta, A., Trombini, M., Labanca, S., Marenco, S., Solarna, D., Pisciotta, L., Baldissarro, I., Picciotto, A., Dellepiane, S. "Hepatic elastometry and Glissonian Line in the assessment of liver fibrosis." Ultrasound in Medicine & Biology 47.4 (2021): 947-959.
https://doi.org/10.1016/j.ultrasmedbio.2020.12.015

- Vestito, L.*, Trombini, M.*, Mori, L., Dellepiane, S., Trompetto, C., Morando, M., Bandini, F. "Improved visuospatial neglect after tDCS and computer-assisted cognitive training in Posterior Cortical Atrophy: a single-case study." Neurocase (2020), 1-7.
https://doi.org/10.1080/13554794.2020.1862242

- Trombini, M., Borro, P., Ziola, S., Dellepiane, S. “A Digital Image Processing Approach for Hepatic Diseases Staging based on the Glisson’s Capsule.” 2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE), Kuala Lumpur, Malaysia, 2020, pp. 1-6.
https://doi.org/10.1109/ICECIE50279.2020.9309633

- Trombini, M., Vestito, L., Morando, M., Mori, M., Trompetto, C., Bandini, F., Dellepiane, S. “Unilateral spatial neglect rehabilitation supported by a digital solution: two case-studies.” 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2020.
https://doi.org/10.1109/EMBC44109.2020.9175834

- Morando, M., Trombini, M., Dellepiane, S. “Application of SVM for Evaluation of Training Performance in Exergames for Motion Rehabilitation.” Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing. 2019.
https://doi.org/10.1145/3332340.3332342

A.2 Posters and Communications

- Simonini, M., Truffelli, R., Ferraro, F., Trombini, M., Dellepiane, S. “Riflessioni sull’utilizzo in clinica degli indicatori del Test di Albert digitalizzato.” 49 Congresso Nazionale SIMFER *Le Radici del Futuro.*

- Truffelli, R., Simonini, M., Ferraro, F., Trombini, M., Dellepiane, S. “Test di Albert in forma digitale. Indicatori e caratterizzazione di un campione di soggetti sani.” XX Congresso Nazionale SIRN *La Neuroriabilitazione ai Tempi del Covid.*

- Doveri, M., Trombini, M., Ferraro, F., Galli, R., Bargeri, A., Rando, S., Dellepiane, S., Bianchi, G. “Towards Systemic Sclerosis Rehabilitation via Videogames” American Congress of Rehumatology – ACR Convergence 2020.

- Doveri, M., Trombini, M., Ferraro, F., Galli, R., Bargeri, A., Rando, S., Dellepiane, S., Bianchi, G. “Sclerodermia e videogiochi: un possible nuovo approccio riabilitativo” 57° Congresso Nazionale SIR.
- Morando, M., Trombini, M., Dellepiane, S. “Remote Evaluation of Rehabilitative Activity and Disability Assessment via Exergames.” Workshop IRT 2019 – Innovation in Rehabilitation Technologies.

A.3 Awards

- Project **STORMS** - Solution Towards Occupational Rehabilitation for Multiple Sclerosis, awarded for *Premio Innovazione Digitale nella Sclerosi Multipla 2020 di Merck*.

A.4 Certificates and courses

- Digitalisation in Space Research, Technische Universität München, Coursera Certificate. Oct 2021.
- Courses on Big Data Analysis, University of California San Diego, Coursera Certificate. Sep 2020.
- Deep Learning Specialization, deeplearning.ai, Coursera Certificate. Dec 2019.
- First MINOA PhD School Mixed-Integer Nonlinear Optimization meets Data Science, CNR - Istituto di Analisi dei Sistemi Informatici (IASI). Jun 2019.

A.5 Conference and events

- IMIP 2021 - 3rd International Conference on Intelligent Medicine and Image Processing. April 23-26, 2021, Tianjin, China (virtual event due to Covid-19 pandemic).

- MIBEC 2020 - IEEE Malaysian International Biomedical Engineering Conference. November 28, 2020, Kuala Lumpur, Malaysia (virtual event due to Covid-19 pandemic).
- HIMS 2020 - World Congress in Computer Science, Computer Engineering, and Applied Computing, International Conference on Health Informatics and Medical Systems. July 27-30, 2020, Las Vegas, USA (virtual event due to Covid-19 pandemic).
- EMBC 2020 - 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), IEEE. July 20-24, 2020 Montréal, Canada (virtual event due to Covid-19 pandemic).
- ISBI 2020 - International Symposium on Biomedical Imaging, IEEE. April 03-07, 2020 Iowa City, USA (virtual event due to Covid-19 pandemic).
- IRT 2019 - Workshop in Innovation in Rehabilitation Technologies. 5 Marzo 2019 Genova, Italia.

A.6 Scientific collaborations

- Centro di riabilitazione srl; Prof. G. Regesta MD.
- Clinica di neuroriabilitazione - IRCCS Ospedale San Martino; Prof. C. Trompetto MD, Prof. L. Mori MD, L. Vestito, G. Schenone.
- Clinica gastroenterologica - IRCCS Ospedale San Martino; Dr. P. Borro MD, A. Pasta MD, S. Ziola MD.
- Department of Biocybernetics and Biomedical Engineering - AGH University of Science and Technology, Cracow, Poland; Prof. Z. Tabor, D. Gut.

- Department of Chemistry and Industrial Chemistry (DCCI) - Università degli Studi di Genova; Prof. G. Petrillo.
- Department of Electrical, Electronics and Telecommunication Engineering and Naval Architecture (DITEN) - Università degli Studi di Genova; Prof. G. Moser, Prof. S. B. Serpico, Dr. A. De Giorgi, Dr. M. Morando, Dr. D. Solarna, F. Ferraro, G. Iaconi, L. Maggiolo.
- Department of Internal Medicine and Medical Specialties (DIMI) - Università degli Studi di Genova; Prof M. A. Cimmino MD.
- Department of Medico-Surgical Sciences and Biotechnologies (DSBMC) - Università La Sapienza; Prof. A. Currà MD.
- Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, and Maternal and Children's Sciences (DINOGMI) - Università degli Studi di Genova; Prof. C. Trompetto MD, Prof. L. Mori MD, L. Vestito.
- Department of Radiology - Jagiellonian University Medical College, Cracow, Poland; Prof. W. Wojciechowski MD, Dr. I. Kucybala MD, K. Krupa MD.
- Società Italiana per lo Studio dullo Stroke (SISS); Prof. F. Bandini MD, Prof. G. Regesta MD.
- Struttura Complessa di Neurologia - ASL3; Prof. F. Bandini MD.
- Struttura Complessa Recupero e Rieducazione Funzionale - ASL3; M. Simonini MD, R. Truffelli.
- Struttura Complessa di Radiologia afferente al Dipartimento Diagnostica - ASL2; G. Ferrero MD.
- Struttura Complessa di Reumatologia - ASL3; G. Bianchi MD, M. Doveri MD.

A.7 Scientific societies

- ACM - Association for Computing Machinery - member.
- CBEES - Chemical, Biological & Environmental Engineering Society - graduate student member.
- CNIT - Consorzio Nazionale Interuniversitario per le Telecomunicazioni - affiliate member.
- CVPL - Associazione Italiana per la ricerca in Computer Vision, Pattern recognition e machine Learning - graduate student member.
- IEEE - Institute of Electrical and Electronics Engineers - graduate student member.
- IEEE BISP Technical Committee - Bio Imaging and Signal Processing - affiliate member.
- IEEE EMBS - Engineering in Medicine and Biology Society - graduate student member.
- IEEE GRSS - Geoscience and Remote Sensing Society - graduate student member.
- IEEE SPS - Signal Processing Society - graduate student member.

A.8 Didactic activities and supervision

- Assistant and Subject Expert for the course Digital Image Processing (Master of science course in Internet and Multimedia Engineering, Master of science course in Electronic Engineering).
- Assistant and Subject Expert for the course Electrical Communications (Bachelor of science course in Biomedical Engineering).

- Assistant and Subject Expert for the course Digital Processing of Historical-Artistic Images (Master course in Art history and valorization of the cultural heritage).
- Co-supervisor of 18 B.Sc. Theses in Biomedical Engineering.
- Co-supervisor of 2 M.Sc. Thesis in Electronic Engineering.
- Co-supervisor of 4 M.Sc. Theses in Internet and Multimedia Engineering.
- Co-supervisor of 1 M.Sc. Thesis in Bio-Engineering.
- Lecturer at Esaote Academy – June - July 2021.

Appendix B

Other works

In this appendix, the other works developed during the PhD, but not described in the present thesis are introduced.

B.1 Ultrasound imaging and hepatic diseases

The present section briefly discloses the work in [195] and [196]. Due to the need for quick and effective treatments for liver diseases, which are among the most common health problems in the world, staging fibrosis through non-invasive and economic methods has become of great importance. Taking inspiration from diagnostic laparoscopy, used in the past for hepatic diseases, ultrasound images of the liver can be studied focusing on a specific region of the organ where the Glisson's capsule is visible. In ultrasound images, the Glisson's capsule appears in the shape of a line which can be extracted via classical methods in literature. By making use of a combination of standard image processing techniques and CNN approaches, one can give evidence to the idea that a great informative potential relies on smoothness of the Glisson's capsule surface. To this purpose, several classifiers are taken into consideration, which deal with different type of data, namely ultrasound images, binary images depicting the Glisson's line, and features vector extracted from the original image. An example of ultrasound image and extracted line is depicted in Figure B.1

The available dataset has been classified according to two, three, and five classes, based on the Metavir score $F0 - F4$ [197]. More specifically, when considering two classes, the dataset was divided into *healthy subjects/low disease patients* (stages $F0$ and $F1$) and

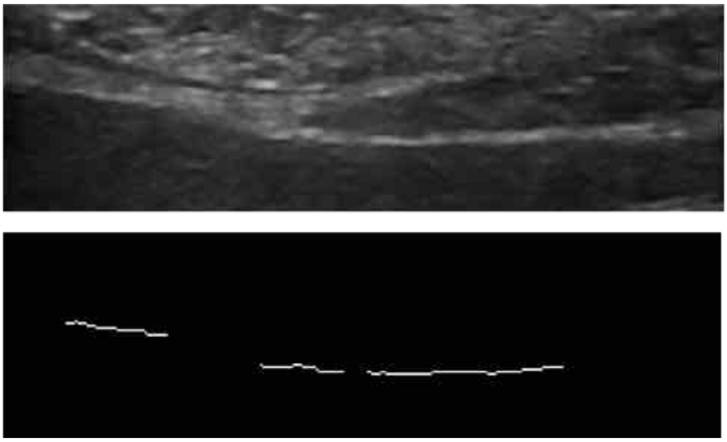


Figure B.1: Example of Glisson's line extraction from a patient's image.

impaired patients (stages $F2$, $F3$, and $F4$); when referring to three classes, the considered groups were *healthy subjects* (stage $F0$), *low/mild level disease patients* (stages $F1$ and $F2$), and *severe stage/cirrhotic patients* (stages $F3$ and $F4$).

The study consisted of four steps: 1) CNN to process only the ultrasound images; 2) CNN to process the ultrasound images together with the extracte Glisson's line; 3) features extraction from the ultrasound images and classification via a shallow neural network; 4) CNN processing of the ultrasound images together with the extracted lines, then concatenated with the output from the shallow network.

The classification accuracy increased as the number of information processed by the model grows. In particular, accuracy increase is mostly significant when features extracted on the basis of the Glisson's line are considered. Then, a small improvement for the classification happens when considering the concatenated model, thus confirming the assumption that the Glisson's capsule is responsible for the highest information content related to the hepatic diseases.

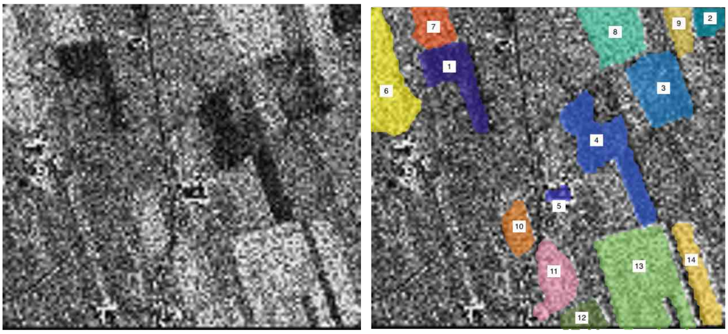
The present study is a preliminary work, due to some limitations i.e. the use of elastometric results as ground truth data rather than the histological examination, and the small sample size (only 157 patients). As for the former issue, ground truth staging was indeed provided by Fibroscan and SWE examination when both the methodology were concordant (to address the problem of user-dependance of SWE diagnosis). However, biopsy is the gold standard for staging liver diseases even though its applicability is limited. Hence, a long time is required to collect enough data to test a classification based on the bioptic exam, which is necessary to validate the proposed approach. In addition, some staging techniques, included Fibroscan and SWE, may depend also on pathological conditions, such as obesity and liver congestion, thus affecting the goodness of the diagnosis. On the contrary, since the proposed study is focused on morphological criteria, based on the Glisson's line, any modifications on the liver status, deriving from pathological conditions, does not influence the results of the current approach.

B.2 Remote sensing applications

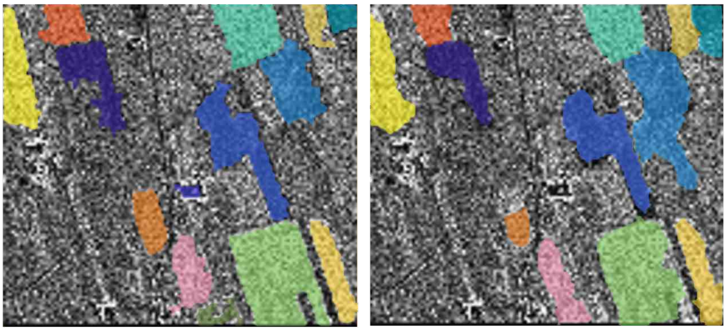
The method presented in 5.1 was also applied to the remote sensing domain. In particular, Synthetic aperture radar (SAR) intensity images were considered. They were acquired either by COSMO-SkyMed in the HH polarization at the spatial resolution of 5m, and depict rural areas in the north of Italy, or by TerraSAR-X in SC mode at the resolution of 18m, acquired in the Mediterranean sea near the French coast.

Concerning the first images, the visible agricultural fields were mainly of three types, corresponding to lower, higher, and intermediate values of the backscattering coefficient and not exhibiting appreciable texture. The proposed segmentation method was applied with the purpose of segmenting the two types of field corresponding to the brighter and darker areas (i.e., higher and lower backscattering), and was also compared with a state-of-the-art method for SAR image segmentation [198].

Figure B.2 shows an example of the application, depicting the original image, the available and manually annotated ground truth data, and the segmentation result obtained via the two aforementioned cases in the considered trials. It is worth mentioning that when referring to the method in [198], ROIs included in the ground truth have been manually selected in the final segmentation map (i.e., by merging oversegmented areas) in order to enable a quantitative comparison.



(a) Data (b) Ground truth



(c) Proposed method (d) Comparison

Figure B.2: Agricultural fields segmentation in SAR imagery.

Similarly to the case of MR images, the use of the MRF model allowed to achieve a spatial regularization in the segmentation map. Indeed, the polygonal shapes of the fields are preserved, despite the speckle [150] that may lead to errors in identifying the borders. In addition, also methods in [26] and [179] were used in this context.

However, they did not result appropriate for such data. In particular, the method in [26] is likely unsuitable for SAR images, were the speckle yields to a significant over-segmentation. Furthermore, the borders of ROIs in SAR data are not well defined, and it is likely the reason why the method in [179] does not perform effectively on such images.

The adopted parametric model for the SAR imagery is the Gamma distribution. Figure B.3 depicts the gray levels histogram of one of the ROIs in the COSMO-SkyMed dataset and the corresponding MoLC-fitted Gamma distribution. The interpretation of the figure is twofold. On one hand, it shows the accuracy of the MoLC estimation for the parameters of the Gamma distribution and, on the other hand, it also visually confirms the Gamma distribution to be a valid model for the statistics of the pixel intensities in the regions detected by the proposed method when applied to this type of image. The same comment holds with regard to the statistics conditioned to other regions in the image.

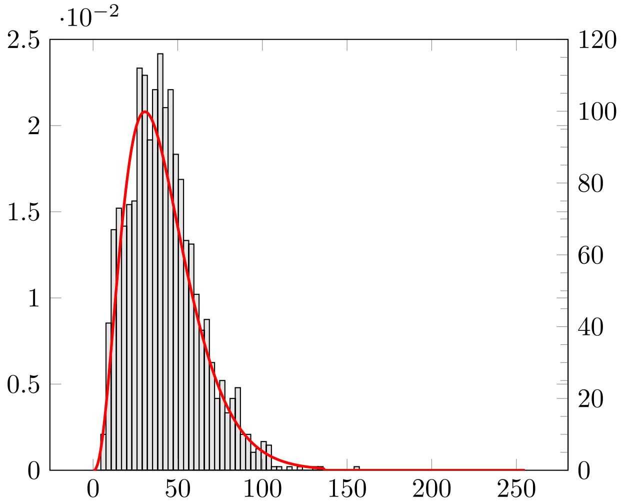


Figure B.3: Gray levels histogram of one of the ROIs in the SAR image depicting the agricultural fields and the MoLC-fitted gamma distribution.

Figure B.4 shows an oil slick located in the Mediterranean sea, along with the segmentation result. The goal is to detect the oil spill, which appears as dark region over the bright background, in order to enable the quantification of its spread. Conversely to the previous case, no ground truth was available in this case study. Yet, qualitative results are provided in order to give a visual feedback on the performances achieved by the proposed method in this new scenario.

Also in this case, the segmentation method was able to achieve satisfactory results. In particular, the segmentation was able to separate the homogeneous regions characterized by the oil spill with the surrounding area, yet the fine details of the oil spill in the upper left part of the images were not identified. This is due to the random grid strategy in the seed placement step. Due to the thin shape of such details, none of the seeds was placed inside the region of interest and, as a consequence, such fine details were not assigned to the homogeneous region associated with the oil spill. The reader may also notice the difference with respect to the output of method [198], where no particular ROI is actually visible.

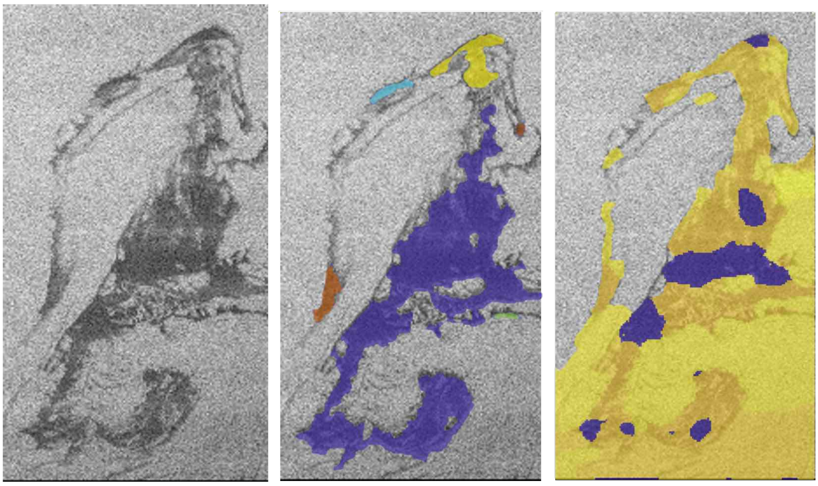

(a) Data (b) Proposed method (c) Comparison

Figure B.4: Oil spill segmentation in SAR images.

B.3 Cultural heritage preservation

In the present section, the work in [199] is briefly described.

Cultural heritage bears witness to life and history, provides an identity to nations, and represents an irreplaceable source of inspiration. Its importance from cultural, historical, and economic points of view is invaluable; thus, its preservation and valorization are crucial topics for our society. Natural aging and deterioration due to external agents endanger artworks such as paintings, sculptures, and architecture, and therefore diagnostic tools are needed for monitoring and preservation.

Monitoring historical artistic heritage consists of the evaluation of possible modifications of some characteristics of the object under observation. When it comes to a artwork or, more generally, a mono- or polychromatic surface, color is one of those characteristics, as it is easily perceivable by the human eye, allows one to distinguish an artwork, and provides information on the nature and status of an artwork.

Color analysis on artworks is generally performed, both qualitatively and quantitatively, via specific instruments such as colorimeters and spectrophotometers, working in a device-independent color space as CIE $L^*a^*b^*$ [200, 201]. This allows for objective assessment of color changes in order to monitor the state of the painting over time and appropriately plan periodic protection and/or restoration actions. Color studies of artworks make also use of Infrared (IR) and Ultraviolet (UV) data (by means of, e.g., infrared reflectography, UV–Visible spectrophotometry, UV reflectance, etc.), or of X-ray fluorescence spectroscopy (XRF) [202, 203].

In the present study, 117 tiles from the database of diagnostic analyses of The Foundation Centre for Conservation and Restoration of cultural heritage La Venaria Reale (in collaboration with the National Institute of Metrological Research and Laboratorio Analisi Scientifiche of Regione Autonoma Valle d'Aosta) are considered

[202]. Figure B.5 shows a table of the considered tiles.

Figure B.5: Picture of a table collecting colored tiles from The Foundation Centre for Conservation and Restoration of cultural heritage La Venaria Reale.

The innovative aspect of the work is the combination of SoA methods with clustered data, in order to have a better correction. In particular, a supervised approach for color correction is proposed, following a cluster analysis based on two different criteria: color and chemical properties of pigments. Aimed at overcoming the issues of small data amount and at finding one-to-one correspondence between image and colorimetric data, samples referring to the same tile are sorted by hue values. This allows for having coupled data and leveraging on supervised methods for a precise and punctual color correction. From the experimental phase, it results that Polynomial Regression [204] is an approach that nicely performs. In addition, the proposed method proved to be effective also with organic pigments, which cannot be analyzed via standard approaches such as XRF.

An example of the obtained result is depicted in Figure B.6. In particular, each visualization depicts the uncalibrated color values, the colorimeter data, and the correction when the method is trained on the whole dataset and on the specific cluster. The reader may notice the improvement on the visual rendering when dealing with clustered data.

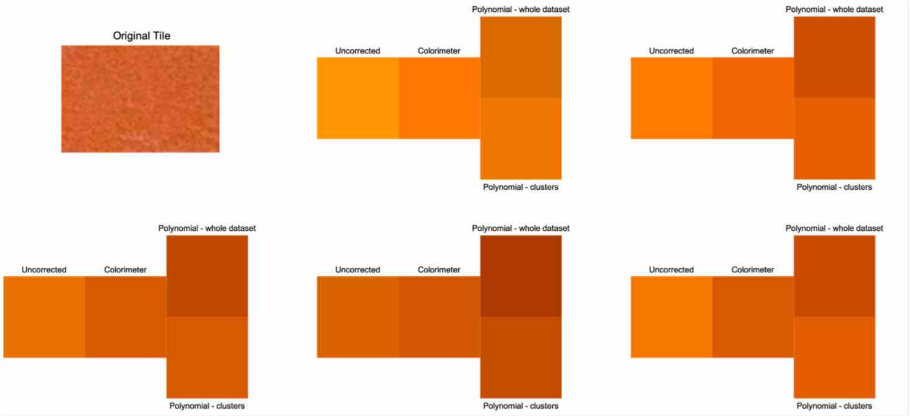


Figure B.6: Figures depicting an example of color correction with respect to the Polynomial Regression method.

B.4 Development cooperation

Private voluntary organizations (PVOs) are major players in responding to many humanitarian needs. For instance, PVOs play a leading role in providing health care and education in the countries where they act. The interest toward the non-profit sector is increasing all around the world, thus denoting a large number of well-intentioned actors. In such a context, the need for coordination between organizations and information access has become relevant, to effectively respond to the major issues for populations to be helped.

In such a framework, an online platform, called SPeRA web-portal, was developed [205]. Its aimed at collecting data related to PVOs operating in Africa. As a starting point, the portal is collecting data related to small Italian PVOs, in order to give them a voice. Indeed, despite their concrete operations and laudable efforts, their activities are barely promoted and described, and that hinders the possibility of joining and channeling efforts, in order to optimize their operations.

Since 2010, the consortium Solidarietà, Progetti e Risorse per l'Africa (SPeRA) is collecting information related to cooperation projects. In order to leverage on such information for driving future projects, based on the necessities that are still not addressed, an online portal describing Italian associations and their activities was created. There are more than 650 projects run by more than 260 different associations. The sectors of activities span from school education to prevention and health care, from support to agricultural and livestock activities to the promotion of entrepreneurial work activities, from the construction of structures and infrastructures to the improvement of the use of local resources. Projects are mostly aimed at helping the local population, with particular attention to women and children, covering many sectors of intervention, in particular prevention and health and many are also projects aimed at helping children and the issue of correct nutrition. Also, the focus is centered towards education, starting from basic school activities, to professional updating for health personnel at various levels, up to cooperation between universities, for example through the creation of an interactive multimedia platform via the web and usable with very low costs and also in situations with a continuous energy supply deficit. To conclude, there are also many interventions in the technological field, aimed at finding solutions to support agriculture, zootechnics, and fishing in many countries, which therefore aim at encouraging the economic development of local communities.

The process of Africa's marginalization is gradually being reversed, and also the perception toward Africa strategic relevance has changed. All of this is highlighting a growing strategic role of Africa in the world economy. For instance, on May 13^{th} 2000, *The Economist* titled "The hopeless continent" by referring to African development. Then, the international sensitivity with respect to Africa started growing in response to African ascending, so that the same newspaper published numbers titled “Africa rising” and “The new scramble for Africa” on December 3^{rd} 2011 and March 9^{th} 2019 respectively. Indeed, democracy is advancing in Africa and that means progress in political stability leading to economic growth. For instance, an important signal of growth comes from the trade agreement African Continental Free Trade Area (ACFTA) [206], which aims at creating a single market fostering the economic advancement of the continent. In addition, it is worth mentioning that African gross domestic product (GDP) was supposed to grow of the 8% each year in average [207], prior the advent of the Covid-19 pandemic which caused a stop.

The goal of the portal is to promote study and research activities for the understanding of the phenomenon of Cooperation in Africa, for statistical analysis, anthropological and geopolitical studies, and of other nature. The final users are the societies themselves, public Institutions, and private companies, who will find a useful tool for the exchange of information and interaction in view of possible joint projects with the aim of improving the effectiveness of the action of the Cooperation.

In general, small PVOs result in being very well integrated in the local society where they operate, but often they are not strictly compliant with the requirements defined by Institutions, as acting independently. Thus, when referring to societies as final user of the portal, it means they can benefit from the exchange of information as it fosters the building of formal contacts with Institutions. Conversely, Institutions can collect and analyze data related to small PVOs, aimed at disseminating the best practice which volunteers should follow.

Finally, a well-defined relation between societies and Institutions can also favor the collaboration with private companies, who can leverage on the experience of PVOs to gather information useful for starting activities addressing the actual needs of the local population.

The importance of the SPeRA portal relies in its possible use for strategic purposes. Accessing data related to humanitarian work in Africa is helpful for defining further operations which address specific needs, which have been identified based on a scientific and precise requirements analysis.

From a technical point of view, the portal deploys PHP, MySql, and SlimFramework v3 technologies for the Back-end, and Ionic, AngularJS, Bootstrap, and Jquery for the Front-end. The Back-end exposes Web Services which, after successful authentication, can be used by the Front-end.

References

[1] D. T. Wade, “Diagnosis in rehabilitation: woolly thinking and resource inequity,” 2002.

[2] P. Suetens, *Fundamentals of medical imaging.* Cambridge university press, 2017.

[3] M. Morando, S. Ponte, E. Ferrara, and S. Dellepiane, “Definition of motion and biophysical indicators for home-based rehabilitation through serious games,” *Information*, vol. 9, no. 5, p. 105, 2018.

[4] H. Azhari, J. A. Kennedy, N. Weiss, and L. Volokh, *From Signals to Image.* Springer, 2020.

[5] J. Haaga, V. Dogra, M. Forsting, R. Gilkeson, K. Hyun, and M. Sundaram, “Ct and mri of the whole body. philadelphia, pa: Mosby,” 2009.

[6] K.-S. Fu and J. Mui, “A survey on image segmentation,” *Pattern recognition*, vol. 13, no. 1, pp. 3–16, 1981.

[7] V. Tomatis, M. A. Cimmino, F. Barbieri, G. Troglio, P. Parascandolo, L. Cesario, G. Viano, L. Vosilla, M. Pitikakis, A. Schiappacasse, *et al.*, “A database of segmented mri images of the wrist and the hand in patients with rheumatic diseases,” in *International Conference on Image Analysis and Processing*, pp. 143–150, Springer, 2015.

[8] J. Włodarczyk, W. Wojciechowski, K. Czaplicka, A. Urbanik, and Z. Tabor, “Fast automated segmentation of wrist bones in magnetic resonance images,” *Computers in biology and medicine*, vol. 65, pp. 44–53, 2015.

[9] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.

[10] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.

[11] S. C. Zhu and A. Yuille, “Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 18, no. 9, pp. 884–900, 1996.

[12] F. Li, M. K. Ng, T. Y. Zeng, and C. Shen, “A multiphase image segmentation method based on fuzzy region competition,” *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 277–299, 2010.

[13] J. Cardinale, G. Paul, and I. F. Sbalzarini, “Discrete region competition for unknown numbers of connected regions,” *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3531–3545, 2012.

[14] D. Jayadevappa, S. Srinivas Kumar, and D. Murty, “Medical image segmentation algorithms using deformable models: a review,” *IETE Technical review*, vol. 28, no. 3, pp. 248–255, 2011.

[15] D. Baltissen, T. Wollmann, M. Gunkel, I. Chung, H. Erfle, K. Rippe, and K. Rohr, “Comparison of segmentation methods for tissue microscopy images of glioblastoma cells,” in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 396–399, IEEE, 2018.

[16] P. Horváth, I. Jermyn, J. Zerubia, and Z. Kato, “A higher-order active contour model for tree detection,” in *18th International Conference on Pattern Recognition (ICPR’06)*, vol. 2, pp. 130–133, IEEE, 2006.

[17] S. Yin, Y. Zhang, and S. Karim, “Large scale remote sensing image segmentation based on fuzzy region competition and gaussian mixture model,” *IEEE Access*, vol. 6, pp. 26069–26080, 2018.

[18] C. A. Bouman and M. Shapiro, “A multiscale random field model for bayesian image segmentation,” *IEEE Transactions on image processing*, vol. 3, no. 2, pp. 162–177, 1994.

[19] Z. Tu and S.-C. Zhu, “Image segmentation by data-driven markov chain monte carlo,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 657–673, 2002.

[20] Z. Tu, X. Chen, A. L. Yuille, and S.-C. Zhu, “Image parsing: Unifying segmentation, detection, and recognition,” *International Journal of computer vision*, vol. 63, no. 2, pp. 113–140, 2005.

[21] Y. Zeng, Y. Zhuge, H. Lu, and L. Zhang, “Joint learning of saliency detection and weakly supervised semantic segmentation,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 7223–7233, 2019.

[22] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, “Salient region detection and segmentation,” in *International conference on computer vision systems*, pp. 66–75, Springer, 2008.

[23] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura, and P. Vandergheynst, “Graph signal processing: Overview, challenges, and applications,” *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[24] L. Gemme and S. Dellepiane, “A new graph-based method for automatic segmentation,” in *International Conference on Image Analysis and Processing*, pp. 601–611, Springer, 2015.

[25] S. Dellepiane and F. Fontana, “Extraction of intensity connectedness for image processing,” *Pattern Recognition Letters*, vol. 16, no. 3, pp. 313–324, 1995.

[26] A. X. Falcão, J. Stolfi, and R. de Alencar Lotufo, “The image foresting transform: Theory, algorithms, and applications,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 1, pp. 19–29, 2004.

[27] J. Cousty, G. Bertrand, L. Najman, and M. Couprie, "Watershed cuts: Minimum spanning forests and the drop of water principle," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 8, pp. 1362–1374, 2008.

[28] J. Cousty and L. Najman, "Incremental algorithm for hierarchical minimum spanning forests and saliency of watershed cuts," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*, pp. 272–283, Springer, 2011.

[29] J. E. Vargas-Muñoz, A. S. Chowdhury, E. B. Alexandre, F. L. Galvão, P. A. V. Miranda, and A. X. Falcão, "An iterative spanning forest framework for superpixel segmentation," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3477–3489, 2019.

[30] X. Ren and J. Malik, "Learning a classification model for segmentation," in *null*, p. 10, IEEE, 2003.

[31] D. Stutz, A. Hermans, and B. Leibe, "Superpixels: An evaluation of the state-of-the-art," *Computer Vision and Image Understanding*, vol. 166, pp. 1–27, 2018.

[32] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[33] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in *Advances in neural information processing systems*, pp. 3320–3328, 2014.

[34] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[35] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[36] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper

with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[37] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[38] N. Ayache and J. Duncan, "20th anniversary of the medical image analysis journal (media)," *Medical Image Analysis*, vol. 33, pp. 1–3, 2016.

[39] W. Chen, W. Li, X. Dong, and J. Pei, "A review of biological image analysis," *Current Bioinformatics*, vol. 13, no. 4, pp. 337–343, 2018.

[40] K. Price, "Anything you can do, i can do better (no you can't)…," *Computer Vision, Graphics, and Image Processing*, vol. 36, no. 2-3, pp. 387–391, 1986.

[41] L. Maier-Hein, M. Eisenmann, A. Reinke, S. Onogur, M. Stankovic, P. Scholz, T. Arbel, H. Bogunovic, A. P. Bradley, A. Carass, *et al.*, "Why rankings of biomedical image analysis competitions should be interpreted with care," *Nature communications*, vol. 9, no. 1, pp. 1–13, 2018.

[42] M. D. Kohli, R. M. Summers, and J. R. Geis, "Medical image data and datasets in the era of machine learning—whitepaper from the 2016 c-mimi meeting dataset session," *Journal of digital imaging*, vol. 30, no. 4, pp. 392–399, 2017.

[43] S. DuBravac and C. Ratti, "The internet of things: evolution or revolution?," *AIG White Paper, http://www. aig. com/Chartis/internet/US/en/AIG% 20White% 20Paper*, 2015.

[44] "Internet of things in healthcare market size, share & trends analysis report by component (service, system & software), by connectivity technology (satellite, cellular), by end use (cro, hospital & clinic), by application, and segment forecasts, 2019 - 2025," accessed 16 September 2020. https://www.grandviewresearch.com/industry-analysis/internet-of-things-iot-healthcare-market/, 2019.

[45] "Iot in healthcare market worth $534.3 billion by 2025 | cagr: 19.9 %," accessed 16 September 2020. https://www.grandviewresearch.com/press-release/global-iot-in-healthcare-market/, 2019.

[46] "Iot in healthcare market by component (medical device, systems & software, services, and connectivity technology), application (telemedicine, connected imaging, and inpatient monitoring), end user, and region - global forecast to 2025," accessed 16 September 2020. https://www.marketsandmarkets.com/Market-Reports/iot-healthcare-market-160082804.html/, 2020.

[47] "World population prospects: The 2019 revision," 2019. New York, United Nations Population Division.

[48] "The global burden of chronic," accessed 16 September 2020. https://www.who.int/nutrition/topics/2_background/en/, 2018.

[49] F. Gabbrielli, L. Bertinato, G. De Filippis, M. Bonomini, and M. Cipolla, "Indicazioni ad interim per servizi assistenziali di telemedicina durante l'emergenza sanitaria covid-19," *Versione del*, vol. 13, 2020.

[50] I. Y. Chung, M. Jung, S. B. Lee, J. W. Lee, Y. R. Park, D. Cho, H. Chung, S. Youn, Y. H. Min, H. J. Park, *et al.*, "An assessment of physical activity data collected via a smartphone app and a smart band in breast cancer survivors: observational study," *Journal of medical Internet research*, vol. 21, no. 9, p. e13463, 2019.

[51] A. Carnevale, U. G. Longo, E. Schena, C. Massaroni, D. L. Presti, A. Berton, V. Candela, and V. Denaro, "Wearable systems for shoulder kinematics assessment: A systematic review," *BMC Musculoskeletal Disorders*, vol. 20, no. 1, p. 546, 2019.

[52] M. F. Domingues, A. Nepomuceno, C. Tavares, A. Radwan, N. Alberto, C. Marques, J. Rodriguez, P. André, and P. Antunes, "Energy-aware wearable e-health architecture using op-

tical fbg sensors for knee kinematic monitoring," in *2018 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2018.

[53] J. R. Saura, "Using data sciences in digital marketing: Framework, methods, and performance metrics," *Journal of Innovation & Knowledge*, 2020.

[54] E. Stanmore, B. Stubbs, D. Vancampfort, E. D. de Bruin, and J. Firth, "The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials," *Neuroscience & Biobehavioral Reviews*, vol. 78, pp. 34–43, 2017.

[55] K. L. Lanctôt, M. P. Lindsay, E. E. Smith, D. J. Sahlas, N. Foley, G. Gubitz, M. Austin, K. Ball, S. Bhogal, T. Blake, *et al.*, "Canadian stroke best practice recommendations: mood, cognition and fatigue following stroke, update 2019," *International Journal of Stroke*, vol. 15, no. 6, pp. 668–688, 2020.

[56] D. Johnson, S. Deterding, K.-A. Kuhn, A. Staneva, S. Stoyanov, and L. Hides, "Gamification for health and wellbeing: A systematic review of the literature," *Internet interventions*, vol. 6, pp. 89–106, 2016.

[57] P. Pullen and W. Seffens, "Machine learning gesture analysis of yoga for exergame development," *IET Cyber-Physical Systems: Theory & Applications*, vol. 3, no. 2, pp. 106–110, 2018.

[58] N. Duarte, O. Postolache, and J. Scharcanski, "Ksgphysiokinect serious game for physiotherapy," in *2014 International Conference and Exposition on Electrical and Power Engineering (EPE)*, pp. 606–611, IEEE, 2014.

[59] D. J. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series.," in *KDD workshop*, vol. 10, pp. 359–370, Seattle, WA, USA:, 1994.

[60] L. Rabiner and B. Juang, "An introduction to hidden markov models," *ieee assp magazine*, vol. 3, no. 1, pp. 4–16, 1986.

[61] R. H. Osgouei, D. Soulsbv, and F. Bello, “An objective evaluation method for rehabilitation exergames,” in *2018 IEEE Games, Entertainment, Media Conference (GEM)*, pp. 28–34, IEEE, 2018.

[62] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[63] J. Weston and C. Watkins, “Multi-class support vector machines,” tech. rep., Citeseer, 1998.

[64] J.-Y. Wang and H.-M. Lee, “Recognition of human actions using motion capture data and support vector machine,” in *2009 WRI World Congress on Software Engineering*, vol. 1, pp. 234–238, IEEE, 2009.

[65] B. Dehbandi, A. Barachant, D. Harary, J. D. Long, K. Z. Tsagaris, S. J. Bumanlag, V. He, and D. Putrino, “Using data from the microsoft kinect 2 to quantify upper limb behavior: a feasibility study,” *IEEE journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1386–1392, 2016.

[66] M. Morando, M. Trombini, and S. Dellepiane, “Application of svm for evaluation of training performance in exergames for motion rehabilitation,” in *Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing*, pp. 1–5, 2019.

[67] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[68] F. Ferraro, M. Trombini, M. Morando, M. Doveri, G. Bianchi, and S. Dellepiane, “Exergames for systemic sclerosis rehabilitation: A pilot study,” in *Advances in Computer Vision and Computational Biology*, pp. 281–291, Springer, 2021.

[69] C. Kwoh, L. Anderson, J. Greene, D. Johnson, J. O’Dell, and M. Robbins, “American college of rheumatology subcommittee on rheumatoid arthritis guidelines. guidelines for the management of rheumatoid arthritis: 2002 update,” *Arthritis Rheum*, vol. 46, pp. 328–346, 2002.

[70] D. Resnick, “Early abnormalities of pisiform and triquetrum in rheumatoid arthritis.,” *Annals of the rheumatic diseases*, vol. 35, no. 1, pp. 46–50, 1976.

[71] J. Włodarczyk, K. Czaplicka, Z. Tabor, W. Wojciechowski, and A. Urbanik, “Segmentation of bones in magnetic resonance images of the wrist,” *International journal of computer assisted radiology and surgery*, vol. 10, no. 4, pp. 419–431, 2015.

[72] A. Fleming, J. M. Crown, and M. Corbett, “Early rheumatoid disease. i. onset.,” *Annals of the Rheumatic Diseases*, vol. 35, no. 4, pp. 357–360, 1976.

[73] M. Østergaard, C. Peterfy, P. Conaghan, F. McQueen, P. Bird, B. Ejbjerg, R. Shnier, P. O’Connor, M. Klarlund, P. Emery, *et al.*, “Omeract rheumatoid arthritis magnetic resonance imaging studies. core set of mri acquisitions, joint pathology definitions, and the omeract ra-mri scoring system.,” *The Journal of rheumatology*, vol. 30, no. 6, pp. 1385–1386, 2003.

[74] W. P. Maksymowych, “Biomarkers for diagnosis of axial spondyloarthritis, disease activity, prognosis, and prediction of response to therapy,” *Frontiers in immunology*, vol. 10, p. 305, 2019.

[75] J. Sieper, M. Rudwaleit, X. Baraliakos, J. Brandt, J. Braun, R. Burgos-Vargas, M. Dougados, K. Hermann, R. Landewe, W. Maksymowych, *et al.*, “The assessment of spondyloarthritis international society (asas) handbook: a guide to assess spondyloarthritis,” *Annals of the rheumatic diseases*, vol. 68, no. Suppl 2, pp. ii1–ii44, 2009.

[76] I. Kucybała, Z. Tabor, J. Polak, A. Urbanik, and W. Wojciechowski, “The semi-automated algorithm for the detection of bone marrow oedema lesions in patients with axial spondyloarthritis,” *Rheumatology international*, vol. 40, no. 4, pp. 625–633, 2020.

[77] R. G. Lambert, P. A. Bakker, D. van der Heijde, U. Weber, M. Rudwaleit, K.-G. A. Hermann, J. Sieper, X. Baraliakos, A. Bennett, J. Braun, *et al.*, “Defining active sacroiliitis on

mri for classification of axial spondyloarthritis: update by the asas mri working group," *Annals of the rheumatic diseases*, vol. 75, no. 11, pp. 1958–1963, 2016.

[78] R. van den Berg, G. Lenczner, F. Thévenin, P. Claudepierre, A. Feydy, M. Reijnierse, A. Saraux, A. Rahmouni, M. Dougados, and D. van der Heijde, "Classification of axial spa based on positive imaging (radiographs and/or mri of the sacroiliac joints) by local rheumatologists or radiologists versus central trained readers in the desir cohort," *Annals of the rheumatic diseases*, vol. 74, no. 11, pp. 2016–2021, 2015.

[79] L. Kwakkenbos, T. A. Sanchez, K. A. Turner, L. Mouthon, M.-E. Carrier, M. Hudson, C. H. van den Ende, A. A. Schouffoer, J. J. Welling, M. Sauvé, *et al.*, "The association of sociodemographic and disease variables with hand function: a scleroderma patient-centered intervention network cohort study," 2018.

[80] A. Del Rosso, S. Maddali Bongi, F. Sigismondi, I. Miniati, F. Bandinelli, and M. Matucci-Cerinic, "The italian version of the hand mobility in scleroderma (hamis) test: evidence for its validity and reliability," *Clinical and experimental rheumatology*, vol. 28, no. 5, p. S42, 2010.

[81] S. Parisi, C. Celletti, M. Scarati, M. Priora, A. Laganà, C. Peroni, F. Camerota, G. La Torre, D. Blow, and E. Fusaro, "Neuromuscular taping enhances hand function in patients with systemic sclerosis: a pilot study," *La Clinica Terapeutica*, vol. 168, no. 6, pp. e371–e375, 2017.

[82] A. Young, R. Namas, C. Dodge, and D. Khanna, "Hand impairment in systemic sclerosis: various manifestations and currently available treatment," *Current treatment options in rheumatology*, vol. 2, no. 3, pp. 252–269, 2016.

[83] N. Mugii, Y. Hamaguchi, and S. Maddali-Bongi, "Clinical significance and usefulness of rehabilitation for systemic sclerosis," *Journal of Scleroderma and Related Disorders*, vol. 3, no. 1, pp. 71–80, 2018.

[84] L. J. Buxbaum, M. Ferraro, T. Veramonti, A. Farne, J. Whyte, E. Ladavas, F. Frassinetti, and H. Coslett, "Hemispatial neglect: Subtypes, neuroanatomy, and disability," *Neurology*, vol. 62, no. 5, pp. 749–756, 2004.

[85] J. Ringman, J. Saver, R. Woolson, W. Clarke, and H. Adams, "Frequency, risk factors, anatomy, and course of unilateral neglect in an acute stroke cohort," *Neurology*, vol. 63, no. 3, pp. 468–474, 2004.

[86] M. Mancuso, S. Rosadoni, D. Capitani, W. Bickerton, G. Humphreys, A. De Tanti, M. Zampolini, G. Galardi, M. Caputo, S. De Pellegrin, *et al.*, "Italian standardization of the apples cancellation test," *Neurological Sciences*, vol. 36, no. 7, pp. 1233–1240, 2015.

[87] M. L. Albert, "A simple test of visual neglect.," *Neurology*, 1973.

[88] M. E. Klinke, T. B. Hafsteinsdóttir, H. Hjaltason, and H. Jónsdóttir, "Ward-based interventions for patients with hemispatial neglect in stroke rehabilitation: a systematic literature review," *International journal of nursing studies*, vol. 52, no. 8, pp. 1375–1403, 2015.

[89] A. Pollock, C. Hazelton, C. A. Henderson, J. Angilley, B. Dhillon, P. Langhorne, K. Livingstone, F. A. Munro, H. Orr, F. J. Rowe, *et al.*, "Interventions for visual field defects in patients with stroke," *Cochrane Database of Systematic Reviews*, no. 10, 2011.

[90] P. Azouvi, S. Jacquin-Courtois, and J. Luauté, "Rehabilitation of unilateral neglect: Evidence-based medicine," *Annals of physical and rehabilitation medicine*, vol. 60, no. 3, pp. 191–197, 2017.

[91] M. A. Nitsche and W. Paulus, "Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation," *The Journal of physiology*, vol. 527, no. 3, pp. 633–639, 2000.

[92] J. Veldema, K. Bösl, G. Neumann, G. Verheyden, and D. A. Nowak, "Noninvasive brain stimulation in rehabilitation of hemispatial neglect after stroke," *CNS spectrums*, pp. 1–12, 2019.

[93] F. Ferraro, M. Trombini, R. Truffelli, M. Simonini, and S. Dellepiane, "On the assessment of unilateral spatial neglect via digital tests," in *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 802–806, IEEE, 2021.

[94] M. Trombini, L. Vestito, M. Morando, L. Mori, C. Trompetto, F. Bandini, and S. Dellepiane, "Unilateral spatial neglect rehabilitation supported by a digital solution: two case-studies," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3670–3675, IEEE, 2020.

[95] L. Vestito, M. Trombini, L. Mori, S. Dellepiane, C. Trompetto, M. Morando, and F. Bandini, "Improved visuospatial neglect after tdcs and computer-assisted cognitive training in posterior cortical atrophy: a single-case study," *Neurocase*, pp. 1–7, 2020.

[96] E. Pedroli, S. Serino, P. Cipresso, F. Pallavicini, I. A. C. Giglioli, E. Guastafierro, A. Gaggioli, and G. Riva, "Neglect app. usability of a new application for assessment and rehabilitation of neglect," in *Proceedings of the 3rd 2015 Workshop on ICTs for improving patients rehabilitation research techniques*, pp. 139–143, ACM, 2015.

[97] G. D'Amico, L. Landucci, and D. Pezzatini, "Natural interactive system for hemispatial neglect rehabilitation," in *International Conference on Image Analysis and Processing*, pp. 501–508, Springer, 2013.

[98] P. B. Gorelick, "The global burden of stroke: persistent and disabling," *The Lancet Neurology*, vol. 18, no. 5, pp. 417–418, 2019.

[99] G. L. Widener, N. Conley, S. Whiteford, J. Gee, A. Harrell, C. Gibson-Horn, V. Block, and D. D. Allen, "Changes in standing stability with balance-based torso-weighting with cerebellar ataxia: A pilot study," *Physiotherapy Research International*, vol. 25, no. 1, p. e1814, 2020.

[100] W. Ilg and D. Timmann, "Gait ataxia—specific cerebellar influences and their rehabilitation," *Movement Disorders*, vol. 28, no. 11, pp. 1566–1575, 2013.

[101] A. Marquer, G. Barbieri, and D. Pérennou, "The assessment and treatment of postural disorders in cerebellar ataxia: a systematic review," *Annals of physical and rehabilitation medicine*, vol. 57, no. 2, pp. 67–78, 2014.

[102] N. H. Bhanpuri, A. M. Okamura, and A. J. Bastian, "Predicting and correcting ataxia using a model of cerebellar function," *Brain*, vol. 137, no. 7, pp. 1931–1944, 2014.

[103] A. M. Zimmet, N. J. Cowan, and A. J. Bastian, "Patients with cerebellar ataxia do not benefit from limb weights," *The Cerebellum*, vol. 18, no. 1, pp. 128–136, 2019.

[104] M. Foroni, C. Mussi, E. Erzili, and G. Salvioli, "L'anziano fragile e il paziente anziano fragile: problema solo terminologico," *Giornale di Gerontologia*, vol. 54, pp. 255–259, 2006.

[105] L. P. Fried, C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. J. Kop, G. Burke, *et al.*, "Frailty in older adults: evidence for a phenotype," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 56, no. 3, pp. M146–M157, 2001.

[106] P. B. da Silva, F. N. Antunes, P. Graef, F. Cechetti, and A. de Souza Pagnussat, "Strength training associated with task-oriented training to enhance upper-limb motor function in elderly patients with mild impairment after stroke: a randomized controlled trial," *American journal of physical medicine & rehabilitation*, vol. 94, no. 1, pp. 11–19, 2015.

[107] K. I. Erickson, M. W. Voss, R. S. Prakash, C. Basak, A. Szabo, L. Chaddock, J. S. Kim, S. Heo, H. Alves, S. M. White, *et al.*, "Exercise training increases size of hippocampus and improves memory," *Proceedings of the National Academy of Sciences*, vol. 108, no. 7, pp. 3017–3022, 2011.

[108] S. Vaynman, Z. Ying, and F. Gomez-Pinilla, "Hippocampal bdnf mediates the efficacy of exercise on synaptic plasticity and cognition," *European Journal of Neuroscience*, vol. 20, no. 10, pp. 2580–2590, 2004.

[109] H.-G. Cha and D.-W. Oh, "Effects of mirror therapy integrated with task-oriented exercise on the balance function of patients with poststroke hemiparesis: a randomized-controlled pilot trial," *International Journal of Rehabilitation Research*, vol. 39, no. 1, pp. 70–76, 2016.

[110] J. A. Anguera, J. Boccanfuso, J. L. Rintoul, O. Al-Hashimi, F. Faraji, J. Janowich, E. Kong, Y. Larraburo, C. Rolle, E. Johnston, *et al.*, "Video game training enhances cognitive control in older adults," *Nature*, vol. 501, no. 7465, p. 97, 2013.

[111] P. M. Keune, S. Hansen, E. Weber, F. Zapf, J. Habich, J. Muenssinger, S. Wolf, M. Schönenberg, and P. Oschmann, "Exploring resting-state eeg brain oscillatory activity in relation to cognitive functioning in multiple sclerosis," *Clinical Neurophysiology*, vol. 128, no. 9, pp. 1746–1754, 2017.

[112] R. H. Benedict, M. P. Amato, J. DeLuca, and J. J. Geurts, "Cognitive impairment in multiple sclerosis: clinical management, mri, and therapeutic avenues," *The Lancet Neurology*, vol. 19, no. 10, pp. 860–871, 2020.

[113] N. D. Chiaravalloti and J. DeLuca, "Cognitive impairment in multiple sclerosis," *The Lancet Neurology*, vol. 7, no. 12, pp. 1139–1151, 2008.

[114] M. Moccia, R. Lanzillo, R. Palladino, K. C.-M. Chang, T. Costabile, C. Russo, A. De Rosa, A. Carotenuto, F. Saccà,

and G. T. Maniscalco, "Cognitive impairment at diagnosis predicts 10-year multiple sclerosis progression," *Multiple Sclerosis Journal*, vol. 22, no. 5, pp. 659–667, 2016.

[115] S. M. Yap, L. O'Donnell, Z. Togher, M. Dillon, N. McNicholas, N. Tubridy, M. Hutchinson, and C. McGuigan, "Safety monitoring of alemtuzumab therapy in active relapsing multiple sclerosis: necessary, manageable but resource-intensive.(p5. 401)," 2017.

[116] G. W. Rebok, K. Ball, L. T. Guey, R. N. Jones, H.-Y. Kim, J. W. King, M. Marsiske, J. N. Morris, S. L. Tennstedt, F. W. Unverzagt, *et al.*, "Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults," *Journal of the American Geriatrics Society*, vol. 62, no. 1, pp. 16–24, 2014.

[117] A. S. Pollock, B. R. Durward, P. J. Rowe, and J. P. Paul, "What is balance?," *Clinical rehabilitation*, vol. 14, no. 4, pp. 402–406, 2000.

[118] S. Verma, D. Kumar, A. Kumawat, A. Dutta, and U. Lahiri, "A low-cost adaptive balance training platform for stroke patients: a usability study," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 7, pp. 935–944, 2017.

[119] G. Saposnik, R. Teasell, M. Mamdani, J. Hall, W. McIlroy, D. Cheung, K. E. Thorpe, L. G. Cohen, and M. Bayley, "Effectiveness of virtual reality using wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle," *Stroke*, vol. 41, no. 7, pp. 1477–1484, 2010.

[120] O. Hornykiewicz, "The discovery of dopamine deficiency in the parkinsonian brain," *Parkinson's Disease and Related Disorders*, pp. 9–15, 2006.

[121] E. Tolosa and C. Pont-Sunyer, "Progress in defining the premotor phase of parkinson's disease," *Journal of the neurological sciences*, vol. 310, no. 1-2, pp. 4–8, 2011.

[122] G.-M. Hariz and L. Forsgren, “Activities of daily living and quality of life in persons with newly diagnosed parkinson’s disease according to subtype of disease, and in comparison to healthy controls,” *Acta Neurologica Scandinavica*, vol. 123, no. 1, pp. 20–27, 2011.

[123] R. Cuevas-Trisan, “Balance problems and fall risks in the elderly,” *Physical Medicine and Rehabilitation Clinics*, vol. 28, no. 4, pp. 727–737, 2017.

[124] Z. Rahmati, S. Behzadipour, A. C. Schouten, and G. Taghizadeh, “A postural control model to assess the improvement of balance rehabilitation in parkinson’s disease,” in *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, pp. 1019–1024, IEEE, 2018.

[125] C. Bayon, A. Emmens, M. Afschrift, T. Van Wouwe, A. Keemink, H. Van Der Kooij, and E. Van Asseldonk, “Can momentum-based control predict human balance recovery strategies?,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 28, no. 9, pp. 2015–2024, 2020.

[126] T. P. Huryn, J.-S. Blouin, E. A. Croft, M. S. Koehle, and H. M. Van der Loos, “Experimental performance evaluation of human balance control models,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 6, pp. 1115–1127, 2014.

[127] C. Franco, A. Fleury, B. Diot, and N. Vuillerme, “Applying entropy to human center of foot pressure data to assess attention investment in balance control,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5586–5589, IEEE, 2018.

[128] T. Bao, B. N. Klatt, S. L. Whitney, K. H. Sienko, and J. Wiens, “Automatically evaluating balance: a machine learning approach,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 27, no. 2, pp. 179–186, 2019.

[129] V. S. Aguilar, J. J. van de Gronde, C. J. Lamoth, N. M. Maurits, and J. B. Roerdink, "Assessing dynamic balance performance during exergaming based on speed and curvature of body movements," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 1, pp. 171–180, 2017.

[130] D. Yoo, Y. Son, D.-H. Kim, K.-H. Seo, and B.-C. Lee, "Technology-assisted ankle rehabilitation improves balance and gait performance in stroke survivors: a randomized controlled study with 1-month follow-up," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 12, pp. 2315–2323, 2018.

[131] S. Gaardsmoe, M. Ovando, K. Bui, and M. J. Johnson, "Development of a low-cost balance assessment system for use in an affordable robot gym in low and middle income countries," in *2020 IEEE 11th Latin American Symposium on Circuits & Systems (LASCAS)*, pp. 1–6, IEEE, 2020.

[132] J. An, J. Kim, E. C. Lai, and B.-C. Lee, "Effects of a smartphone-based wearable telerehabilitation system for in-home dynamic weight-shifting balance exercises by individuals with parkinson's disease," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 5678–5681, IEEE, 2020.

[133] A. Fung, E. C. Lai, and B.-C. Lee, "Usability and validation of the smarter balance system: An unsupervised dynamic balance exercises system for individuals with parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 4, pp. 798–806, 2018.

[134] M. A. Eltoukhy, C. Kuenze, J. Oh, and J. F. Signorile, "Validation of static and dynamic balance assessment using microsoft kinect for young and elderly populations," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 147–153, 2017.

[135] A. Mengarelli, S. Cardarelli, A. Strazza, F. Di Nardo, S. Fioretti, and F. Verdini, "Validity of the nintendo wii bal-

ance board for the assessment of balance measures in the functional reach test," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 26, no. 7, pp. 1400–1406, 2018.

[136] B. L. Luu, T. P. Huryn, H. M. Van der Loos, E. A. Croft, and J.-S. Blouin, "Validation of a robotic balance system for investigations in the control of human standing balance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 4, pp. 382–390, 2011.

[137] F. Noveletto, A. Soares, B. Mello, C. Sevegnani, F. Eichinger, M. d. S. Hounsell, and P. Bertemes-Filho, "Biomedical serious game system for balance rehabilitation of hemiparetic stroke patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 11, pp. 2179–2188, 2018.

[138] M. Trombini, F. Ferraro, M. Morando, G. Regesta, and S. Dellepiane, "A solution for the remote care of frail elderly individuals via exergames," *Sensors*, vol. 21, no. 8, p. 2719, 2021.

[139] Z. Kato, J. Zerubia, *et al.*, *Markov random fields in image segmentation.* Nov, 2012.

[140] J. Besag, J., "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological)* 36.

[141] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[142] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[143] H. Greenspan, A. Ruf, and J. Goldberger, "Constrained gaussian mixture model framework for automatic segmentation of mr brain images," *IEEE transactions on medical imaging*, vol. 25, no. 9, pp. 1233–1245, 2006.

[144] Z. Ji, Y. Xia, Q. Sun, Q. Chen, D. Xia, and D. D. Feng, "Fuzzy local gaussian mixture model for brain mr image segmentation," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 339–347, 2012.

[145] D.-X. Yue, F. Xu, A. C. Frery, and Y.-Q. Jin, "Synthetic aperture radar image statistical modeling: Part one-single-pixel statistical models," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 1, pp. 82–114, 2020.

[146] D. Solarna, G. Moser, and S. B. Serpico, "A markovian approach to unsupervised change detection with multiresolution and multimodality sar data," *Remote Sensing*, vol. 10, no. 11, p. 1671, 2018.

[147] D. Solarna, G. Moser, and S. B. Serpico, "Multiresolution and multimodality sar data fusion based on markov and conditional random fields for unsupervised change detection," in *2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 29–32, IEEE, 2019.

[148] H.-C. Li, V. A. Krylov, P.-Z. Fan, J. Zerubia, and W. J. Emery, "Unsupervised learning of generalized gamma mixture model with application in statistical modeling of high-resolution sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2153–2170, 2015.

[149] H.-C. Li, W. Hong, and Y.-R. Wu, "Generalized gamma distribution with molc estimation for statistical modeling of sar images," in *2007 1st Asian and Pacific Conference on Synthetic Aperture Radar*, pp. 525–528, IEEE, 2007.

[150] F. T. Ulaby, R. K. Moore, and A. K. Fung, "Microwave remote sensing: Active and passive. volume 3-from theory to applications," 1986.

[151] A. N. Gavrilov and I. M. Parnum, "Fluctuations of seafloor backscatter data from multibeam sonar systems," *IEEE Journal of Oceanic Engineering*, vol. 35, no. 2, pp. 209–219, 2010.

[152] P. Sudeep, P. Palanisamy, J. Rajan, H. Baradaran, L. Saba, A. Gupta, and J. S. Suri, "Speckle reduction in medical ultrasound images using an unbiased non-local means method," *Biomedical Signal Processing and Control*, vol. 28, pp. 1–8, 2016.

[153] V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "On the method of logarithmic cumulants for parametric probability density function estimation," *IEEE Transactions on Image Processing*, vol. 22, no. 10, pp. 3791–3806, 2013.

[154] J.-M. Nicolas and S. N. Anfinsen, "Introduction to second kind statistics: Application of log-moments and log-cumulants to the analysis of radar image distributions," *Trait. Signal*, vol. 19, no. 3, pp. 139–167, 2002.

[155] M. Abramowitz and I. Stegun, "Handbook of mathematical functions: With formulas, graphs, and mathematical tables," *National Bureau of Standards Applied Mathematics Series. New York, NY, USA: Dover*, 1964.

[156] M. Hazewinkel, "Newton method," *Encyclopedia of Mathematics, Springer*, 2001.

[157] J. Chen, Z. Li, and B. Huang, "Linear spectral clustering superpixel," *IEEE Transactions on image processing*, vol. 26, no. 7, pp. 3317–3330, 2017.

[158] J. Yao, M. Boben, S. Fidler, and R. Urtasun, "Real-time coarse-to-fine topologically preserving segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2947–2955, 2015.

[159] B. Jähne, *Digital Image Processing 6th Edition.* Springer.

[160] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE transactions on pattern analysis and machine intelligence*, no. 4, pp. 532–550, 1987.

[161] E. Ferrara, S. Ponte, M. Morando, and S. Dellepiane, "Removes remote monitoring validation engineering system: New

way of care," in *Italian Forum of Ambient Assisted Living*, pp. 227–240, Springer, 2017.

[162] M. Pasha and S. M. W. Shah, "Framework for e-health systems in iot-based environments," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.

[163] Y. Ma, Y. Wang, J. Yang, Y. Miao, and W. Li, "Big health application system based on health internet of things and big data," *IEEE Access*, vol. 5, pp. 7885–7897, 2016.

[164] K. Ullah, M. A. Shah, and S. Zhang, "Effective ways to use internet of things in the field of medical and smart health care," in *2016 International Conference on Intelligent Systems Engineering (ICISE)*, pp. 372–379, IEEE, 2016.

[165] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer, "Assessment and calibration of a rgb-d camera (kinect v2 sensor) towards a potential use for close-range 3d modeling," *Remote Sensing*, vol. 7, no. 10, pp. 13070–13097, 2015.

[166] D. J. Geerse, B. H. Coolen, and M. Roerdink, "Kinematic validation of a multi-kinect v2 instrumented 10-meter walkway for quantitative gait assessments," *PloS one*, vol. 10, no. 10, p. e0139913, 2015.

[167] S. T. Pöhlmann, E. F. Harkness, C. J. Taylor, and S. M. Astley, "Evaluation of kinect 3d sensor for healthcare imaging," *Journal of medical and biological engineering*, vol. 36, no. 6, pp. 857–870, 2016.

[168] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A. U. Brandt, and T. Schmitz-Hübsch, "Accuracy and reliability of the kinect version 2 for clinical measurement of motor function," *PloS one*, vol. 11, no. 11, p. e0166532, 2016.

[169] L. H. Larsen, L. Schou, H. H. Lund, and H. Langberg, "The physical effect of exergames in healthy elderly—a systematic review," *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, vol. 2, no. 4, pp. 205–212, 2013.

[170] K. M. Gerling, J. Schild, and M. Masuch, “Exergame design for elderly users: the case study of silverbalance,” in *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology*, pp. 66–69, ACM, 2010.

[171] B. Wilson, J. Cockburn, and P. Halligan, “Development of a behavioral test of visuospatial neglect.,” *Archives of physical medicine and rehabilitation*, vol. 68, no. 2, p. 98, 1987.

[172] N. Cristianini, J. Shawe-Taylor, *et al.*, *An introduction to support vector machines and other kernel-based learning methods.* Cambridge university press, 2000.

[173] C. Olah, “Understanding lstm networks.”

[174] M. Nguyen, “Illustrated guide to lstm’s and gru’s: A step by step explanation.”

[175] L. Gemme and S. G. Dellepiane, “An automatic data-driven method for sar image segmentation in sea surface analysis,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 5, pp. 2633–2646, 2018.

[176] S. Z. Li, *Markov Random Field Modeling in Image Analysis.* Advances in Computer Vision and Pattern Recognition, Springer-Verlag London, 2009.

[177] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “Slic superpixels compared to state-of-the-art superpixel methods,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[178] H.-H. Chang, A. H. Zhuang, D. J. Valentino, and W.-C. Chu, “Performance measure characterization for evaluating neuroimage segmentation algorithms,” *Neuroimage*, vol. 47, no. 1, pp. 122–135, 2009.

[179] T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[180] K. Rzecki, I. Kucybała, D. Gut, A. Jarosz, T. Nabagło, Z. Tabor, and W. Wojciechowski, "Fully automated algorithm for the detection of bone marrow oedema lesions in patients with axial spondyloarthritis–feasibility study," *Biocybernetics and Biomedical Engineering*, 2021.

[181] P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to data mining.* Pearson Education India, 2016.

[182] M. Trombini, F. Ferraro, and S. Dellepiane, "On the use of boundary gradient for the analysis of mr wrist bones volumes segmentation," *2021 3rd International Conference on Intelligent Medicine and Image Processing (IMIP)*, In press.

[183] L. Shapiro and G. Stockman, "Computer Vision," 2000.

[184] H. Zhang, J. E. Fritts, and S. A. Goldman, "Entropy-based objective evaluation method for image segmentation," in *Storage and Retrieval Methods and Applications for Multimedia 2004*, vol. 5307, pp. 38–49, Intern. Society for Optics and Photonics, 2003.

[185] M. Trombini, F. Ferraro, A. Nardelli, L. Vestito, G. Schenone, L. Mori, C. Trompetto, and S. Dellepiane, "On the performance assessment during the practice of an exergame for cerebellar ataxia patients," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 1–5, IEEE, 2021.

[186] M. D. Lezak, D. B. Howieson, D. W. Loring, J. S. Fischer, *et al.*, *Neuropsychological assessment.* Oxford University Press, USA, 2004.

[187] G. Gainotti, P. D'Erme, and P. Bartolomeo, "Early orientation of attention toward the half space ipsilateral to the lesion in patients with unilateral brain damage.," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 54, no. 12, pp. 1082–1089, 1991.

[188] S. J. Crutch, M. Lehmann, J. M. Schott, G. D. Rabinovici, M. N. Rossor, and N. C. Fox, "Posterior cortical atrophy," *The Lancet Neurology*, vol. 11, no. 2, pp. 170–178, 2012.

[189] S. C. Beh, B. Muthusamy, P. Calabresi, J. Hart, D. Zee, V. Patel, and E. Frohman, “Hiding in plain sight: a closer look at posterior cortical atrophy,” *Practical neurology*, vol. 15, no. 1, pp. 5–13, 2015.

[190] T. Schmitz-Hübsch, S. T. Du Montcel, L. Baliko, J. Berciano, S. Boesch, C. Depondt, P. Giunti, C. Globas, J. Infante, J.-S. Kang, *et al.*, “Scale for the assessment and rating of ataxia: development of a new clinical scale,” *Neurology*, vol. 66, no. 11, pp. 1717–1720, 2006.

[191] L. E. Powell and A. M. Myers, “The activities-specific balance confidence (abc) scale,” *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 50, no. 1, pp. M28–M34, 1995.

[192] D. Leightley and M. H. Yap, “Digital analysis of sit-to-stand in masters athletes, healthy old people, and young adults using a depth sensor,” in *Healthcare*, vol. 6, p. 21, Multidisciplinary Digital Publishing Institute, 2018.

[193] R. Ganea, A. Paraschiv-Ionescu, C. Büla, S. Rochat, and K. Aminian, “Multi-parametric evaluation of sit-to-stand and stand-to-sit transitions in elderly people,” *Medical engineering & physics*, vol. 33, no. 9, pp. 1086–1093, 2011.

[194] A. Papoulis and H. Saunders, “Probability, random variables and stochastic processes,” 1989.

[195] M. Trombini, P. Borro, S. Ziola, and S. Dellepiane, “A digital image processing approach for hepatic diseases staging based on the glisson’s capsule,” in *2020 2nd International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, pp. 1–6, IEEE, 2020.

[196] P. Borro, S. Ziola, A. Pasta, M. Trombini, S. Labanca, S. Marenco, D. Solarna, L. Pisciotta, I. Baldissarro, A. Picciotto, *et al.*, “Hepatic elastometry and glissonian line in the assessment of liver fibrosis,” *Ultrasound in Medicine & Biology*, vol. 47, no. 4, pp. 947–959, 2021.

[197] G. Shiha and K. Zalata, "Ishak versus metavir: terminology, convertibility and correlation with laboratory changes in chronic hepatitis c," in *Liver Biopsy*, IntechOpen, 2011.

[198] R. Shang, J. Lin, L. Jiao, and Y. Li, "Sar image segmentation using region smoothing and label correction," *Remote Sensing*, vol. 12, no. 5, p. 803, 2020.

[199] M. Trombini, F. Ferraro, E. Manfredi, G. Petrillo, and S. Dellepiane, "Camera color correction for cultural heritage preservation based on clustered data," *Journal of Imaging*, vol. 7, no. 7, p. 115, 2021.

[200] N. A. Ibraheem, M. M. Hasan, R. Z. Khan, and P. K. Mishra, "Understanding color models: a review," *ARPN Journal of science and technology*, vol. 2, no. 3, pp. 265–275, 2012.

[201] P. Sanmartín, E. Chorro, D. Vázquez-Nion, F. M. Martínez-Verdú, and B. Prieto, "Conversion of a digital camera into a non-contact colorimeter for use in stone cultural heritage: The application case to spanish granites," *Measurement*, vol. 56, pp. 194–202, 2014.

[202] T. Cavaleri, P. Buscaglia, S. Migliorini, M. Nervo, G. Piccablotto, A. Piccirillo, M. Pisani, D. Puglisi, D. Vaudan, and M. Zucco, "Pictorial materials database: 1200 combinations of pigments, dyes, binders and varnishes designed as a tool for heritage science and conservation," *Applied Physics A*, vol. 123, no. 6, p. 419, 2017.

[203] A. Pelagotti, A. Del Mastio, A. De Rosa, and A. Piva, "Multispectral imaging of paintings," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 27–36, 2008.

[204] G. Hong, M. R. Luo, and P. A. Rhodes, "A study of digital camera colorimetric characterization based on polynomial modeling," *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish*

Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, vol. 26, no. 1, pp. 76–84, 2001.

[205] "Spera consortium (solidarietà, progetti e risorse per l'africa i.e., solidarity, projects and resources for africa)," accessed 31 May 2021. https://www.consorziospera.org/portale-italiaafrica.

[206] "Acfta african continental free trade area," accessed 31 May 2021. https://au.int/en/ti/cfta/about.

[207] G. Mistretta, *Africa's pathways: the future of the Continent through Europe, Italy, China and New Actors*. Luiss university press, 2021.