Analysis and optimization of disease monitoring via computational solutions”

Sofia Ostellino

The diagnose of neurological diseases presents a challenge, but disease monitoring does even more so: the evolution of the symptoms and the characteristics of a disease continue to evolve over the course of the patient's life. Here is where the development of computational methods can have a real impact on the quality of clinical care, as the heterogeneity of neurological conditions requires the integration of different types of clinical data. The aim of computational solutions is to solve real clinical needs and answer to questions that are part of the everyday life of doctors: this is why, during the Research here presented, there has always been a special consideration for the opinion of the experts, an attention to the state-of-the-art literature, and a scrupulousness towards solutions that are also flexible and scalable. The aim of the Research was therefore answer to a variety of questions, addressing practical and clinical needs:

- What are the main needs when developing computational solutions for the monitoring of neurological diseases such as MS and Alzheimer’s disease? How straightforward is it to use public datasets and publicly available tools?
- Is it useful to use different types of clinical data for disease monitoring?
- Can the solutions developed for the other inquiries be applied to cover the needs of another field, such as the gene therapy with AAVs?

The first step was an in-depth analysis of the state-of-the-art to gain knowledge in the field. Afterwards, various sources of data were used to identify how the raw data could be processed, modified, or combined, to extract information via Deep Learning. Challenges and questions related to different topics were addressed adapting, according to the needs, the processing pipeline for the images, the methods of data collection, analysis and storage, and the deep learning frameworks trained.

The processing pipeline that was developed fills the gap between heterogenous datasets and the use of these as datasets for the training of Deep Learning networks: its performances are comparable to the state-of-the-art, but it does not rely on external plug-ins, as many of the existing tools, avoiding inconsistencies and versioning issues. Furthermore, it guarantees direct access to functions’ parameters for customizations.

Moreover, the analysis and practical use of the ADNI database for exploring the feasibility of computational solutions for the monitoring of Alzheimer’s Disease, allowed to explore the use of different clinical data for the prediction of disease conversion in time through of a Deep Learning framework. The use of tabular and imaging data, and the choice of specific types of imaging modalities was evaluated in different settings. The approach used to tackle the aforementioned topics was then transferred to the field of gene therapy, on the possibility of improving the analysis of the vectors that carry the genome of interest. The ITR structures, characteristic of every vector, are the focus of the analysis, given their importance and attention in the state-of-the-art: the use of the graphic method of the dotplots gave interesting results.

The methodology followed can be, in retrospect, defined effective: the research questions posed at the beginning lead to a in depth-analysis of the state-of-the-art both from a technical and a medical point of view. The approach required – since the beginning – to acquire hands-on experience on the tools proposed by the literature, and on the sources of data essential to the developing of the proposed solutions. The variety of topics that were subject of research constituted both a resource and a limitation: on one hand, they directed every developing step to be scalable and flexible and, on the other hand, they limited the available time to focus only on one research question/topic.