Feeding Neuroimaging Repositories

Y. Vives-Gilabert [1], J. Delgado-Mengual [1], M. Delfino-Reznicek [1], B. Gómez-Anson [2]

[1]Port d'Informació Científica, Edifici D Campus UAB, Barcelona
[2]Hospital de la Santa Creu i Sant Pau, Barcelona

1. CONTEXT

Thanks to recent technological advances, the data extracted from post-processing medical images is becoming very useful in the research of neurodegenerative diseases. Large image databases are necessary to perform statistically significant studies whose conclusions can be applied to the early diagnosis of such diseases. How to feed these kinds of data repositories is addressed in this paper.

2. HOW TO FEED THE NEUROIMAGING DATA REPOSITORY AT PIC

Two ways of feeding the repository have been implemented:

1) Manually through the PICNIC web portal. In this case, the medical researchers have to download MRIs from the scanner, perform a de-identification and send them to the repository with a pre-established structure of the files through the PICNIC (https://neuroweb01.pic.es) web portal with a java applet called JUPLOAD. One of the disadvantages of this way of transfer the data is that the structure of files has to be build manually by the researcher. If there are errors in the name of the folder or the structure is not correctly built, the transfer will present errors. Moreover, the transfer time will depend on the internet connection of the user, being recommended to transfer 2 or 3 subjects at the same time.

2) Automatically through a "MRI absorber" bridge between the hospital and PICNIC. The architecture of the automatic transfer mode is presented in Fig. 1.

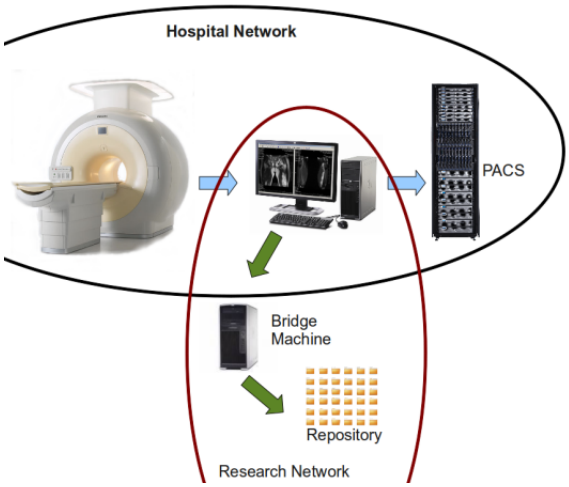

Fig. 1 Architecture of the automatic feeding

- The MRI scanner: it has independent connections to two different networks: the network from the hospital and the research network. The MRI technicians send a set of MRIs (different sequences) from the patients included in research studies using the same application provided by the manufacturer of the scanner to the MRI gateway through the DICOM protocol.

- The "MRI absorber" bridge: this hardware bridge has a DICOM receiver installed, allowing the reception of the images from the scanner. Once a MRI study comes into this machine, it is saved, encrypted, in a temporary folder that will contain the raw data of the project and it will remain there until the end of the project. Afterwards, a de-identification process will start in order to anonymize the header of the MRIs and erase all the personal data, extracting the data related to the gender and age of the patient (useful information for research purposes) before. After that, a parser of the DICOMDIR file starts. The DICOMDIR file contains an index of all the files transmitted through the DICOM protocol, with a relation of the files and the sequence at which they belong to. The objective of the parser is to organize the anonymized data into the structure of the repository, to prepare it for the transmission. Finally, the transmission port of the machine will be activated and the transfer will be performed to PICNIC through an encrypted channel. Once the transfer will be finished, the port will automatically be closed again.

- PICNIC: when the data arrive to the repository, it is saved into the hardware system with the rest of the images. PICNIC system will detect the arrival of new data and it will include the metadata in the database.

3. CONCLUSIONS

In this paper, two ways of feeding neuroimaging repositories have been presented: one manually through a web portal and a second one automatically through a machine bridge between the hospital and the repository.