Development of a Probabilistic Brain Atlas and Tissue Probability Models

X. Beltran Urbano[1] and F. Hartmann[1]

[1] *University of Girona, Erasmus Mundus Joint Master Degree in Medical Imaging and Applications (MAIA)*

Abstract—The accurate segmentation of brain tissues in magnetic resonance imaging (MRI) is a critical component in the analysis of neurological conditions as well as in neuroscientific research. This paper presents a comprehensive pipeline for the development and implementation of a probabilistic brain atlas alongside tissue probability models using a dataset of MRI images. The methodology integrates image registration techniques, including rigid, affine, and non-rigid transformations, to align images with a common space. A probabilistic atlas is then constructed, representing the statistical distribution of tissue types across the dataset. Further, tissue probability models are generated to facilitate the translation of voxel intensity information into tissue-specific probability maps. The resultant atlas and models provide a robust framework for improved brain tissue segmentation, offering detailed anatomical references that can enhance the diagnostic process and contribute to the precision of neuroscientific studies. The validity of our approach is demonstrated through the qualitative evaluation of the atlas and probability models, which show high precision and clear differentiation of brain tissues.

Keywords—Probabilistic Brain Atlas, Tissue Probability Models, Magnetic Resonance Imaging (MRI), Image Registration, Neurological Segmentation, Gaussian Smoothing, Voxel Intensity Mapping

I. INTRODUCTION

Image registration is an essential task in medical imaging as it enables different images to be aligned into a shared spatial frame of reference. This procedure is crucial for a range of applications, including comparing patient scans over time, combining information from diverse imaging modalities, and mapping individual images to standardised anatomical atlases. Through techniques such as rigid, affine, or non-rigid transformations, image registration adjusts for translational, rotational, and scaling discrepancies between images. This allows the integration of multi-temporal or multi-modal datasets, enabling the synthesis of information vital for pathology identification and monitoring, treatment planning, and the improvement of our understanding of brain anatomy and functional connectivity. Following the meticulous process of image registration, the creation of a probabilistic atlas becomes the subsequent pivotal step. This atlas is constructed by aggregating and analysing the aligned images from a diverse population, which allows for the determination of the statistical probability of tissue types at each voxel across the dataset. The main goal of this laboratory is to establish a comprehensive pipeline for creating a probabilistic atlas. This includes registering the dataset images and developing tissue models that assign probabilities to intensity values corresponding to different brain tissues.

II. DATASET

To implement the probabilistic atlas, we have utilised a dataset composed of 15 different MRI images. For each of the cases, we have the following modalities:

- T1-weighted (T1)
- Ground Truth (GT) of the different brain tissues
- Mask of the Brain tissue area

An example of the different image modalities in the dataset can be observed in Figure 1.

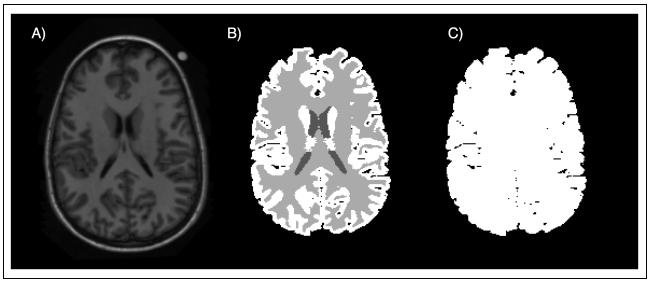

Fig. 1: Example of the different modalities in the dataset. **A)** T1 **B)** Ground Truth **C)** Mask of the Brain Tissue Area

III. METHOD: PROBABILISTIC ATLAS

In this section, the procedures that have been followed to carry out the probabilistic brain atlas and the tissue probability models will be presented.

I. Registration

In order to propagate the labels and build the atlas, the images need to be registered to a fixed image. As the chosen fixed image will have an impact on the resulting atlas, it is beneficial to choose the fixed image as the image most similar to all other images. In order to compute the similarity measure, the

images need to be registered; to be more specific, the images need to have the same dimensions. For this, rigid registration is performed. Afterwards, the similarity measures are computed, and the fixed image is chosen. Finally, the registrations are computed using the fixed image found. The different steps of the mentioned approach can be observed in the Figure 2.

1. Rigid Registration

In order to compute the similarity between the images in the dataset, we first need to register them. To implement this task, we have chosen a rigid transformation due to its simplicity and low computational cost. This transformation is a geometric transformation that preserves distances and angles between points. This means that under a rigid transformation, an object is rotated and translated but not scaled or distorted. In a three-dimensional space, a rigid transformation involves translation and rotation, and it can be represented by the following formula:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_x \\ R_{21} & R_{22} & R_{23} & t_y \\ R_{31} & R_{32} & R_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{1}$$

where:

- (x', y', z') are the coordinates after transformation.
- (x, y, z) are the coordinates before transformation.
- R_{ij} are the elements of the 3x3 rotation matrix R, which represents rotation around the x, y, and z axes.
- (t_x, t_y, t_z) are the translations along the x, y, and z axes.

To perform the registration we used the Elastix library [1].

2. Similarity Measure

In order to align our dataset accurately, we employed the Mean Squared Error (MSE) to measure the similarity between images, averaging this metric over all voxels during each comparison. After computing the MSE of a given image with respect to all the other images, the average of all the MSE values has been computed. Then, the image that yielded the lowest average MSE, which happened to be case 1010, was selected as the fixed image for the final registration. A lower MSE indicates a greater similarity between images, leading to minimal differences and optimal alignment throughout the dataset.

The average mean squared error (MSE) for the i-th image in the dataset with respect to all other images is given by the formula:

$$\overline{\text{MSE}}_i = \frac{1}{M} \sum_{j=1}^{M} \left(\frac{1}{N} \sum_{k=1}^{N} (I_{i,k} - I_{j,k})^2 \right) \tag{2}$$

where:

- M is the total number of images in the dataset.
- N is the number of voxels in each image.
- $I_{i,k}$ is the intensity of the k-th voxel in the i-th image.
- $I_{j,k}$ is the intensity of the k-th voxel in the j-th image.

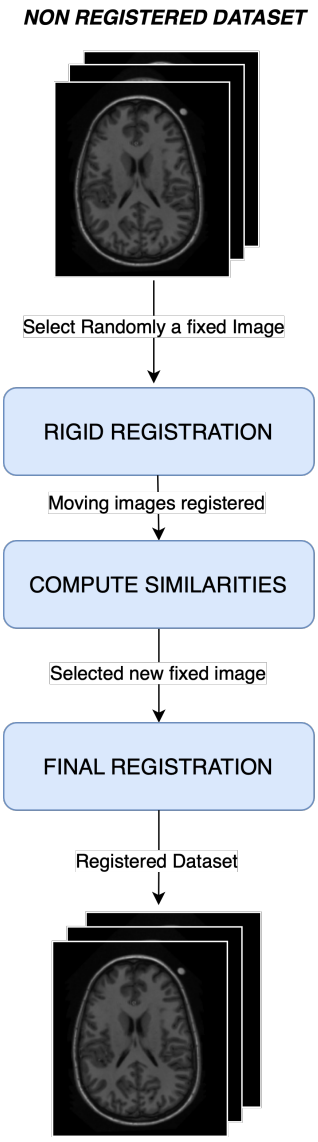

Fig. 2: Workflow of the different steps of the registration approach.

3. Final Registration

Finally, the registrations are recomputed with the previously chosen fixed image. Instead of a rigid registration, a multi-step and multi-resolution registration is employed. First, the images are registered rigidly. Secondly, the images are registered with an affine registration. Thirdly, a b-spline multi-resolution registration is performed. Here, an image pyramid with six levels is chosen. Each level's resolution is half that of the previous level. Furthermore, an advanced normalised correlation metric with a transform bending energy penalty is chosen as proposed by [2]. All parameter files are available onmodelZoo.

II. Label Propagation

To create the probabilistic atlas, we needed to apply the previously computed registration to the labels of each of the images in the dataset. Using the transformation matrices obtained in the previous section, in this step we transformed each of the label images separately. This is done to prevent any overlap or interpolation errors between the labels.

III. Building the Probabilistic Atlas

After the label propagation is carried out, we construct our probabilistic brain atlas by computing the voxel-wise mean across all propagated labels of each image, treating each label distinctly. In the Equation 3, we can observe the formula used to implement this task.

$$P_k(x|k) = \frac{1}{N}\sum_{i=1}^{N} M_{k,i}(x|k) \tag{3}$$

Where:

- $P_k(x)$ is the probability of voxel x belonging to a class k. The sum of the different class probabilities given a specific voxel, will sum 1.
- N is the total number of images in the dataset.
- $M_{k,i}(x|k)$ is the intensity value at voxel x for class k in the i-th image.

IV. Building the mean image

A template or mean image is commonly used for registration or analysis options at a later point. In order to compute this image, the voxel-wise mean of the registered moving images and the fixed image are taken.

$$I(x,y,z) = \frac{1}{N}\sum_{k=1}^{N} I_k(x,y,z) \tag{4}$$

IV. METHOD: TISSUE MODELS

The tissue probabilistic model is a strategy in which the voxel intensity information of each of the images in our dataset is translated into a map of tissue probabilities. This map will provide us with the probability of a given intensity value belonging to each of the tissue labels.

To compute this model, we have followed the next steps:

1. Intensity Distribution

To build the tissue models, we first computed the intensity profiles of our dataset by applying the Ground Truth (GT) masks to each image, acquiring the intensity values, and storing them into separate vectors for each label. This gave us three distinct vectors representing the intensities of white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF). Before doing so, in order to ensure consistent intensity values across different images and to facilitate accurate comparison and analysis, we employ Min-Max normalisation to normalise the intensity values of each image within a range of [0,255]. Finally, we created a histogram from each vector and combined them to form a comprehensive histogram that maps out the distribution of tissue intensities.

2. Histogram of Tissue Probabilities

After computing the histogram, we implemented a histogram of tissue probabilities. To develop this task, we followed the next steps:

- **Normalisation 1:** We normalised the histograms by dividing the values in each class by the total sum of values for that class, thereby scaling the histograms to ensure they all fit within the same range on the y-axis. Additionally, the area of the histogram is now one, transforming the histogram into a probability distribution.
- **Normalisation 2:** We normalised the probability distribution to ensure that the sum of probabilities for each intensity value adds up to one.
- **Gaussian Filter:** Finally, to smooth out sharp peaks and create a more gradual distribution of tissue probabilities, we applied a Gaussian filter with a standard deviation of $\sigma = 20$ to the histograms of each label.

  After smoothing the histogram with the Gaussian filter, we generated a csv file containing three columns that represent the probabilities for each tissue type for every intensity value.

V. RESULTS

In this section, some qualitative results corresponding to the registration, probabilistic atlas, and tissue models will be presented.

To begin, Figure 6 shows an example of the outcomes of our final registration. Then, in Figure 7, an example of several planes (Axial, Coronal, and Sagittal) of distinct slices of our probabilistic atlas constructed in subsection III, is shown. Following that, in Figure 3 and Figure 4, we can see the evolution of our tissue probability model presented in Figure 5. Finally, in Figure 8, we can see an example of our Mean Image, which was developed in subsection IV.

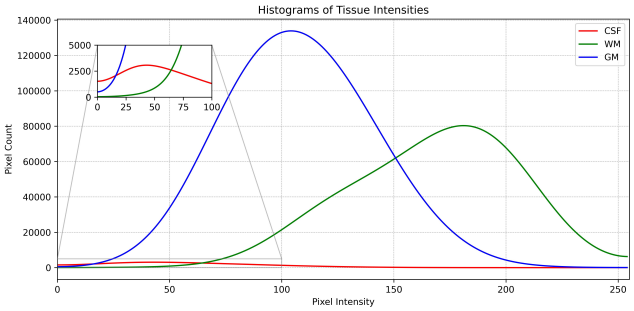

Fig. 3: Intensity distribution of the different tissues in the dataset.

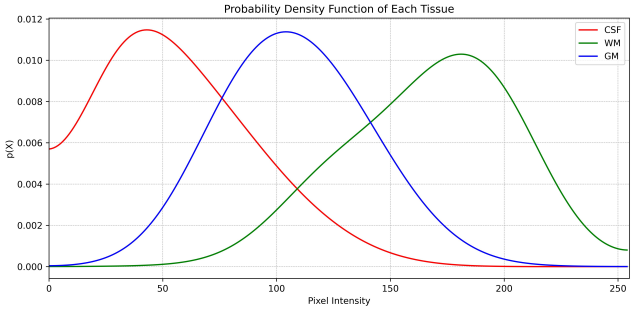

Fig. 4: Normalized intensity distribution of the different tissues in the dataset.

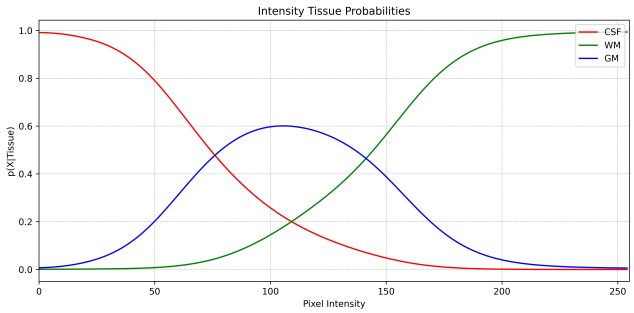

Fig. 5: Tissue probability.

VI. Qualitative Evaluation

As visible from Figure 7, the atlas has probabilities close to one (white) for regions of homogeneous tissue. Furthermore, the probabilities in the boundary regions go smoothly from one to zero (dark grey to light gray). From looking at the probabilistic atlas alone, differentiation between the different tissues is relatively simple, as the expected tissue shapes are apparent.

While the mean image looks slightly blurry in the boundary regions, the different brain structures can be clearly identified. The skull is relatively smooth through the entire image and doesn't show big jumps. In combination, these results should allow for a good registration of the atlas.

Due to the applied Gaussian smoothing, the tissue models look very smooth. Furthermore, the tissues are in the expected range and order in terms of intensities. First, the cerebrospinal fluid. Second is the grey matter, and last is the white matter. After normalising a new image to the intensity range of 0 to 255, the mapping of the tissue probabilities to the intensities should be accomplishable and beneficial for segmentation tasks under the assumption that the new image has a similar distribution.

VII. Conclusion

The present study effectively constructed a comprehensive pipeline to generate probabilistic brain atlas and tissue probability models by capitalising on the robust functionalities of image registration. Our methodology encompassed meticulous steps, starting from rigid registration to label propagation, and culminated in the construction of a finely detailed probabilistic atlas. The resultant probabilistic atlas demonstrated a high level of precision, with clear demarcations of tissue boundaries and homogeneous regions. Furthermore, the tissue probability histograms, normalised and smoothed, offer a robust reference for mapping tissue types in MRI scans, enhancing the potential for accurate diagnostic applications. As we look ahead, we are excited to use the presented project to address brain tissue segmentation. By employing the intensity information provided by the tissue models and the spatial information obtained from the probabilistic atlas, these novel tools will enable us to execute segmentation with greater precision. We anticipate that by doing so, the segmentation's performance will be significantly enhanced, thereby facilitating the resolution of neuroimaging challenges that may assist in the diagnosis, treatment, and follow-up of millions of individuals afflicted with neurological disorders.

VIII. Design and implementation

Throughout the development of this project, we rigorously followed the principles of object-oriented programming and constructed our solution from scratch. The theoretical concepts that were covered in the theory lectures served as the foundation for the entire implementation, guaranteeing a robust foundation based on established academic principles. By following this systematic approach, we were able to acquire a more profound comprehension of every element of the algorithm as we manually pieced together every function in the project classes.

IX. Project Management

This project, which was developed from the ground up by two team members, required innovative input as well as thorough planning. This project was completed in four lab hours, as expected, due to our good basis in the theoretical issues of the tactics used. Our time management strategy has always been largely focused on increasing productivity inside lab hours by leveraging the benefits of pair programming to improve work execution and successfully troubleshoot. The journey from a blank slate to a fully-functional probabilistic brain atlas and tissue probabilistic model demonstrates our dedication to developing this project with the highest quality and efficiency.

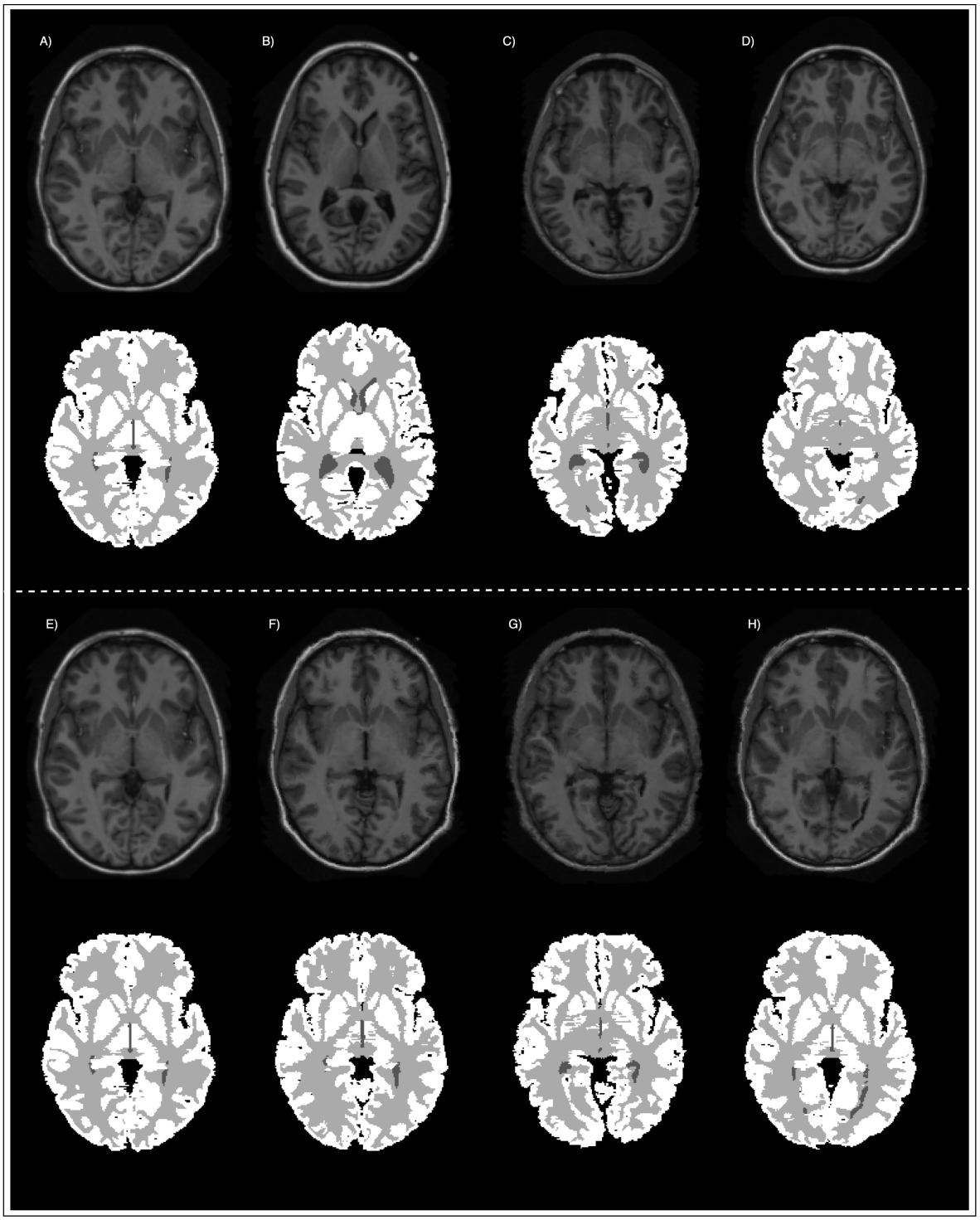

Fig. 6: Example of the registration results. In the first half of the figure,we include the original image and corresponding ground truth. A,B, C, and D correspond to the images 1010 (fixed), 1013, 1036, and 1017, respectively. In the second half of the figure, we can observe the registration of both the original image and the ground truth. E,F, G, and H correspond to the images 1010 (fixed), 1013, 1036, and 1017, respectively.

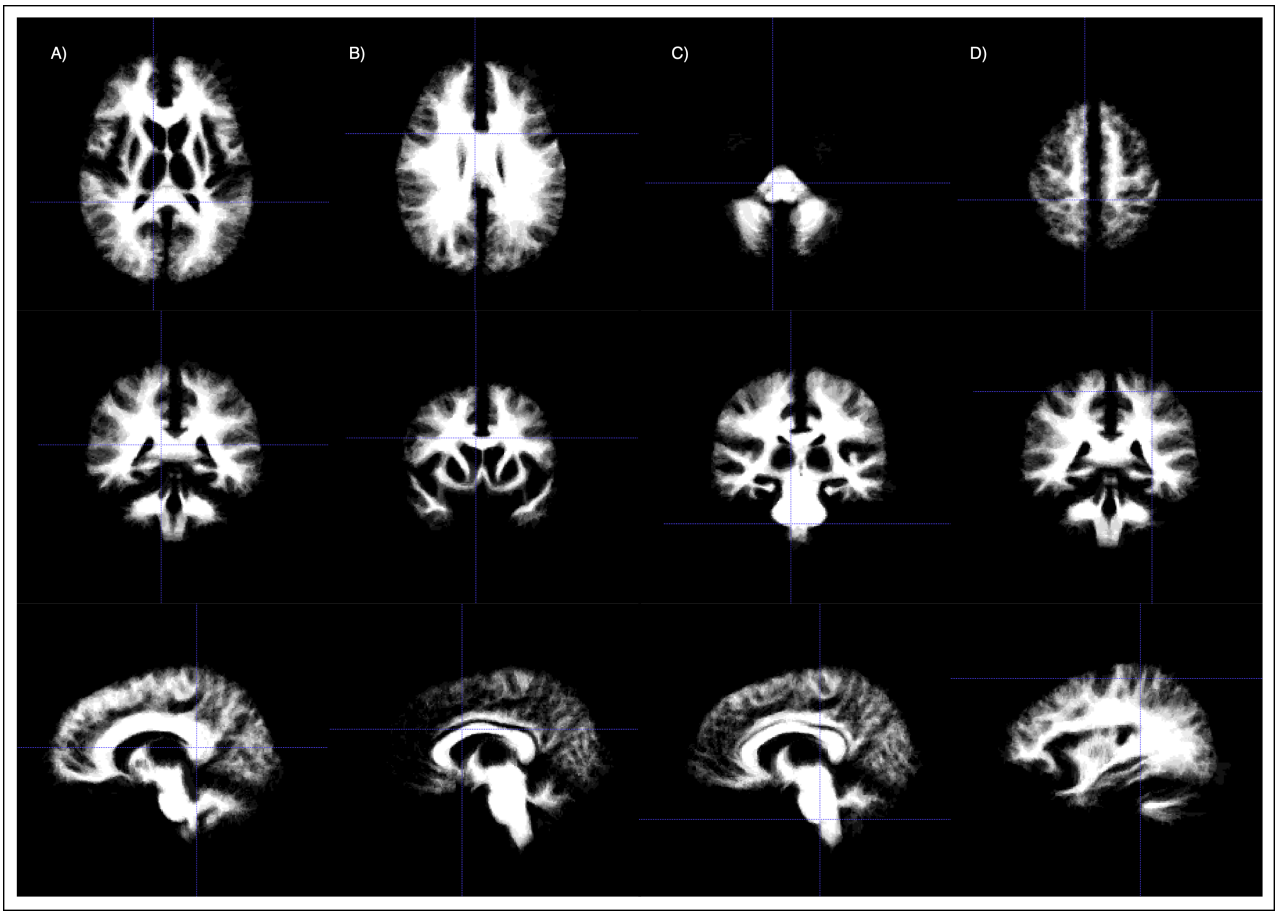

Fig. 7: Example of different planes (Axial, Coronal and Sagital) of different slices of our probabilistic atlas. **A)** Slice nº 145, **B)** Slice nº156, **C)** Slice nº91, **D)** Slice nº185.

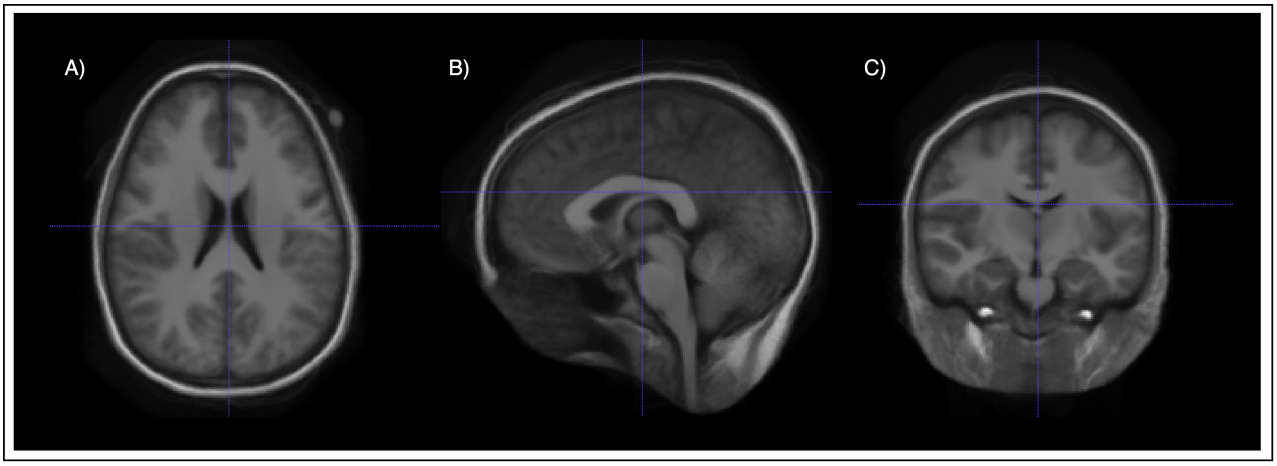

Fig. 8: Example of slice nº151 of our final Mean Image.

REFERENCES

[1] Stefan Klein et al. "elastix: A Toolbox for Intensity-Based Medical Image Registration". In: *IEEE Transactions on Medical Imaging* 29.1 (2010), pp. 196–205. DOI: 10.1109/TMI.2009.2035616.

[2] Inge A. Mulder et al. "Automated Ischemic Lesion Segmentation in MRI Mouse Brain Data after Transient Middle Cerebral Artery Occlusion". In: *Frontiers in Neuroinformatics* 11 (2017). ISSN: 1662-5196. DOI: 10.3389/fninf.2017.00003. URL: https://www.frontiersin.org/articles/10.3389/fninf.2017.00003.