Laboratory Report I
SPM: A Neuroimaging Tool

Joaquin Oscar SEIA
Kaouther MOUHEB

Context

Brain tissue segmentation is a first step in many neuroimaging applications. It can provide useful information to the clinicians and help in the diagnosis and prognosis of multiple brain diseases. Statistical Parametrical Mapping (SPM) is a neuroimaging tool written in MatLab that offers multiple functionalities related to brain image processing, one of which is tissue segmentation.

Tissue segmentation is performed using a probabilistic generative approach, in which by means of the Bayes' rule we obtain the posterior probabilities of each voxel belonging to each of the tissue classes.

$$p(\omega_i|I(x,y,z)) = \frac{p(I(x,y,z)|\omega_i) * p(\omega_i)}{p(I(x,y,z))} \quad (1)$$

The method will determine a posterior probability for each voxels as indicated in Equation 1. This expression involves the following elements:

1. Likelihood $p(I(x,y,z)|\omega_i)$: the probability that given that a voxel (in the normalized space registered image) belong to the tissue class, then the voxel has that particular intensity. According to SPM manual [1] the likelihood is modeled/estimated using a Gaussian mixture model, where more than one Gaussian are combined to model the intensity distribution of all the pixels belonging to one class.

2. Prior probability $p(\omega_i)$: this represents the domain knowledge indicating how probable each tissue class is in each voxel. This information is presented as probability maps for each class which are obtained from atlases, in our case we utilized the *ICBM-European brains* one. This maps results from registering all the brain volumes of the subjects contained in the atlas to a normalized space and using their ground truth labels to determine how probable is that one voxel in this normalized space to belong to each of the classes.

3. Marginalized probability $p(I(x,y,z))$: the probability that a particular intensity occurs in a voxel.

4. Posterior probability $p(\omega_i|I(x,y,z))$: the probability that given that we have a certain intensity in a certain voxel in the normalized registered space, the tissue class is ω_i.

After getting the posterior probability maps for each class, they can be re-warped to the original image space applying the inverse transformation of the previously mentioned registration procedure. To obtain the final segmentation a Bayesian classifier is used, which consists in assigning to each voxel the class with the largest posterior probability, theoretically this guarantees the minimum possible classification error.

Objectives

The aim of this work is to inspect the SPM tool, understand the algorithms behind it, be able to modify their parameters and use it to successfully perform brain tissue segmentation. The complete code developed for this labwork can be found at: https://github.com/kakou34/misa_lab.

Methodology

To achieve the objectives of this course work the following process was followed

1. The SPM tool and the ITK Snap visualization software were installed. Five brain MRI images were downloaded to test the different algorithms.

2. The ITK Snap tool was used to inspect the given data examples.

3. SPM was used programmatically to perform bias field correction with different parameters trying find an optimal combination for them that let us obtain the best segmentation.

4. The segmentation posterior probability maps of the different regions were generated by SPM using both single and multichannel approaches. In the first case, T1 and T2 FLAIR sequences were used separately and in the former they were used

	Regularization	CSF	GM	WM	All Tissues
FWHM = 60	0.0	0.7709 ∓ 0.0339	0.7487 ∓ 0.0302	0.8129 ∓ 0.0315	0.7775 ∓ 0.0404
	0.0001	**0.7721** ∓ 0.0315	**0.7506** ∓ 0.0294	**0.8149** ∓ 0.0317	**0.7792** ∓ 0.0398
	0.01	0.7667 ∓ 0.0317	0.7487 ∓ 0.0312	0.8148 ∓ 0.0313	0.7767 ∓ 0.0410
	1.0	0.7661 ∓ 0.0328	0.7475 ∓ 0.0321	0.8143 ∓ 0.0314	0.7759 ∓ 0.0416

Table 1: Mean dice scores across patients for each tissue type - T1 images.

	Reg.	S1	S2	S3	S4	S5
FWHM = 60	0.0	0.7924 ± 0.036	0.7401 ± 0.028	0.7776 ± 0.046	**0.7951** ± 0.027	0.7822 ± 0.060
	0.0001	**0.7939** ± 0.037	**0.7428** ± 0.028	**0.7791** ± 0.046	0.7948 ± 0.026	**0.7854** ± 0.059
	0.01	0.7912 ± 0.037	0.7400 ± 0.030	0.7769 ± 0.047	0.7920 ± 0.026	0.7835 ± 0.062
	1.0	0.7906 ± 0.037	0.7384 ± 0.031	0.7765 ± 0.047	0.7918 ± 0.026	0.7823 ± 0.063

Table 2: Mean dice scores across tissue type for each subject - T1 images.

jointly (as recommended in SPM's manual) to obtain a better segmentation.

5. The skull stripping step was performed based on the generated segmentation masks.

6. The brain tissue segmentation results were evaluated both qualitatively and quantitatively using the provided ground truth labels of each volume.

The following sections will further explain and illustrate each of the steps of the process.

Bias Field Correction

Bias field is an artifact seen in magnetic resonance images as a low-frequency, smooth, spatially varying signal [2].

The SPM tool allows to modify two parameters related to the bias field, the **regularisation parameter** and the **average resolution FWHM** (full width at half maximum).

The *regularization parameter* is used to include prior knowledge about the distribution of the fields likely to be encountered by the correction algorithm during the optimization process. This allows the algorithm to consider the differences between smooth intensity changes due to the MR physics and the high frequency details that come from the different brain tissues. With this in mind, if the data is known to have no artifacts this regularisation parameter should be set to a high value to penalize the algorithm if the bias field is assigned high parameters.[1]

On the other hand, the *FWHM* is a parameter to model the Gaussian smoothness of the bias field. It should be set to a high value if prior knowledge is available supporting the presence of a very smooth intensity non-uniformity artifact. [1]

To inspect the effect of these parameters two experiments were done. In the first one, while setting the FHWM to 60 mm cutoff, the following values of regularization were tested: 0 (no regularisation), 0.0001 (very light regularisation), 0.01 (medium regularization), 1 (very heavy regularization). In the second one, the FWHM was set to different values (40mm, 60mm, 80mm, 100mm, 120mm cutoffs) while maintaining the regularisation parameter at 0.01.

T1, T2 FLAIR sequences were tested individually and also combined in a multichannel approach. The resulting preprocessed images were saved together with the bias field by setting the write parameter of the SPM segmentation function to 1.

Results

Example results of slice 24 from the first sample are shown in figure 1 for T1 and figure 2 for T2 FLAIR. The figures show the original slice, the same slice after bias field correction, and the corrected bias field respectively. Examples from the remaining samples can be found in Appendix I.

In the case of T1 weighted scans, tables 1 and 2 show that across tissues and across subjects the regularization of 0.0001 is a good choice for the parameter. This tells us, that in all acquisitions the magnetic field was quite non-homogeneous, even more in subject 4, since a low penalization means higher bias field is obtained.

Tables 3 and 4 show that across tissues and across subjects the FWHM of 40mm is a good choice for the parameter. This again confirms us that in all acquisitions the magnetic field was quite non-homogeneous where lower smoothing allows for more local variations of the magnetic field.

In a similar way, the best bias correction hyperparameters were obtained for T2 FLAIR scans (the complete tables are omitted to avoid unnecessary increasing the report's extension). When setting the FMHW to 60mm a regularization of 1 resulted in the

	FWHM	CSF	GM	WM	All Tissues
	40	**0.7676** ± 0.0316	**0.7488** ± 0.0314	0.8142 ± 0.0316	**0.7769** ± 0.0400
Reg =	60	0.7667 ± 0.0317	0.7487 ± 0.0312	0.8148 ± 0.0313	0.7767 ± 0.0410
0.001	80	0.7667 ± 0.0318	0.7484 ± 0.0313	**0.8151** ± 0.0313	0.7767 ± 0.0412
	100	0.7666 ± 0.0321	0.7481 ± 0.0316	0.8150 ± 0.0314	0.7766 ± 0.0414
	120	0.7663 ± 0.0326	0.7479 ± 0.0317	0.8147 ± 0.0313	0.7763 ± 0.0415

Table 3: Mean dice scores across patients for each tissue type - T1 images.

	FWHM	S1	S2	S3	S4	S5
	40	**0.7915** ± 0.036	0.7386 ± 0.029	**0.7779** ± 0.046	**0.7923** ± 0.026	**0.7839** ± 0.061
Reg =	60	0.7912 ± 0.036	**0.7400** ± 0.029	0.7769 ± 0.047	0.7920 ± 0.026	0.7835 ± 0.062
0.001	80	0.7912 ± 0.037	0.7399 ± 0.030	0.7771 ± 0.047	0.7919 ± 0.026	0.7836 ± 0.063
	100	0.7912 ± 0.037	0.7393 ± 0.030	0.7772 ± 0.047	0.7917 ± 0.026	0.7833 ± 0.063
	120	0.7909 ± 0.037	0.7391 ± 0.030	0.7771 ± 0.047	0.7916 ± 0.026	0.7827 ± 0.063

Table 4: Mean dice scores across tissue type for each subject - T1 images.

highest mean dice score both subject and tissue class wise. Fixing the regularization at 0.01, the FWHM of 120mm maximized the mean dice score across subject and tissue type. This let us infer that the bias field in the T2-FLAIR image was rather homogeneous. Finally, the multichannel (T1+T2-FALIR) approach supported by SPM was tried. In the same manner as before, regularization equal to 0.01 and FWHM of 100 where found to produce the best results across tissues and patients.

According to these results[1] the following experiments were done using the combination of best regularization and FMHW parameters found for each modality.

Skull Stripping

Skull stripping is a preliminary step used in many MRI analysis sequences before the application of other algorithms. It consists of removing the bones and extra-meningeal tissues from the brain image. Although there is no direct way to obtain the skull-stripped version of an MRI brain image in SPM, this process can be achieved by segmenting the anatomical/structural scan, and masking the original brain volume by the estimated Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) masks.

The following steps were used for skull stripping:

1. First, the anatomical scan of the brain was segmented. The SPM tool generated six probability maps: CSF, GM, WM, Skull, Meninges and Air (background).

2. Each voxel in the brain volume was assigned to the class of the highest posterior probability.

3. The brain mask is generated by assigning the voxels belonging to one of three classes WM, GM or CSF a value of 1 and setting the remaining voxels to a value of 0.

4. Multiplying the original volume (T1 or T2-FLAIR sequences) element-wise with this binary mask provides the skull-stripped version of the scan.

The following code snippet implements these steps:

```
% generate segmentation using SPM
result = spm_seg(structural_fn, ...
    settings);
% Read tissue probability maps
gm = niftiread(result.gm_fn(1:end-2));
wm = niftiread(result.wm_fn(1:end-2));
csf = niftiread(result.csf_fn(1:end-2));
bone = ...
    niftiread(result.bone_fn(1:end-2));
soft = ...
    niftiread(result.soft_fn(1:end-2));
air = niftiread(result.air_fn(1:end-2));

% Classify the tissues according to ...
    maximum a posteriori probabilies
maps = zeros(240, 240, 48, 6);
maps(:, :, :, 1) = csf;
maps(:, :, :, 2) = gm;
maps(:, :, :, 3) = wm;
maps(:, :, :, 4) = bone;
maps(:, :, :, 5) = soft;
maps(:, :, :, 6) = air;
[¬, res_seg] = max(maps, [], 4);

% Extract brain mask
brain_mask = int16(res_seg<4); % ...
    Keep GM+WM+CSF
skull_mask = int16(res_seg==4); % ...
    bone mask

% remove sull
skull_stripped = vol .* brain_mask;
```

[1] Ideally a grid search should have been performed to analyse the synergy between the parameters, the computational burden prevented us from doing it.

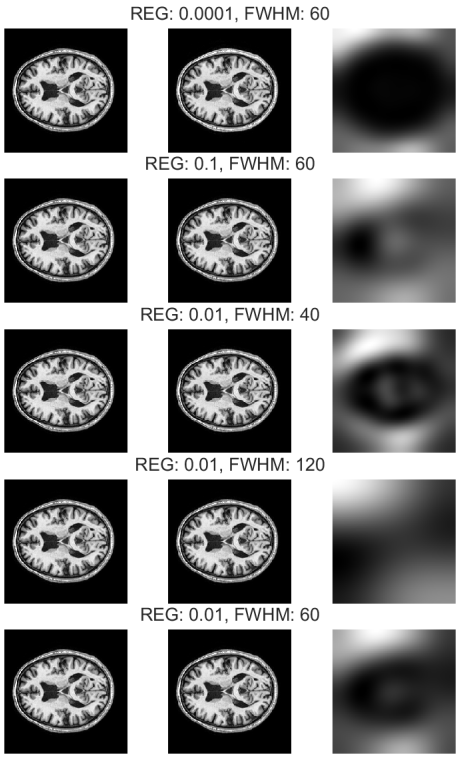

Figure 1: Example bias field correction results - Sample 1 - T1 sequence (slice 24)

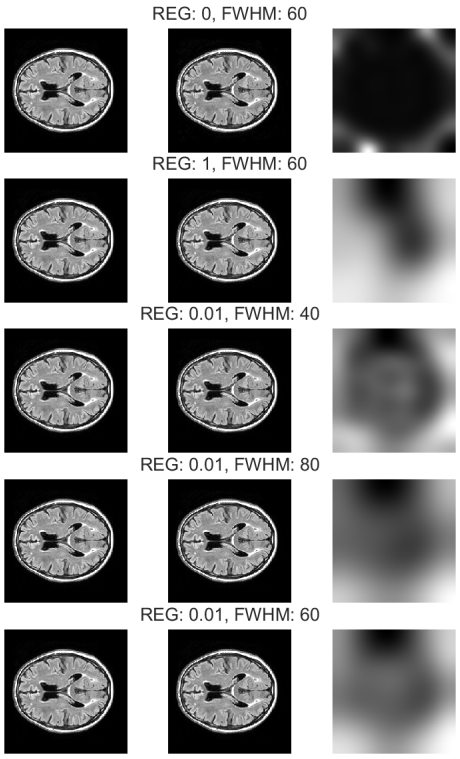

Figure 2: Example bias field correction results - Sample 1 - T2 FLAIR sequence (slice 24)

Results

Figure 3a shows the results of the skull stripping step for the first sample using T1 sequence, figure 3b shows the results using the T2 FLAIR sequence while figure 3c shows the results obtained using the multichannel approach (Note that the skull stripping is shown on the original T1 sequence). The results for the remaining subjects can be found in Appendix I.

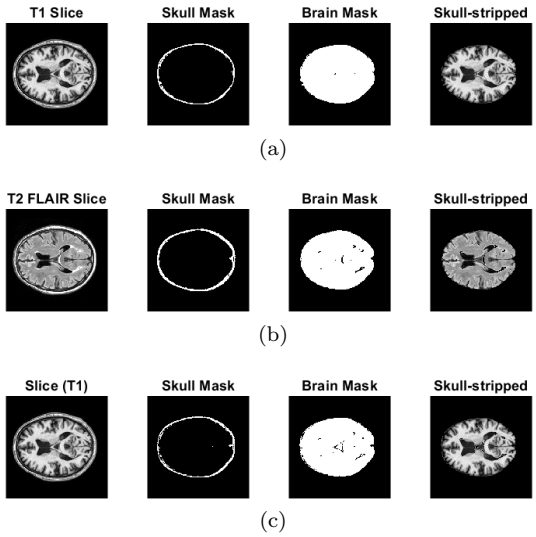

Figure 3: Skull stripping results for the first subject obtained using (a) T1 (b) T2 FLAIR (c) T1 + T2 FLAIR sequence images (slice 24).

A quantitative evaluation of the skull stripping performance was performed by calculating the dice score for each patient and for each modality between the generated brain mask and the brain mask obtained from the ground truth labels (CSF+GM+WM masks).

$$DS = \frac{2TP}{FN + 2TP + FP} \tag{2}$$

Where TP denotes voxels correctly classified, FN voxels classified as background but belonging to the foreground class, and FP voxels classified as belonging to the foreground class being from the background.

We obtained a dice score of **0.9129** $\pm$ **0.0114** across all patients using T1 modality, **0.9140** $\pm$ **0.0075** using T2 FLAIR and **0.9131** $\pm$ **0.0064** using both modalities together.

Tissue Segmentation

Tissue segmentation was performed programmatically using the function *spm_seg()* in order to avoid the usage of SPM's graphical user interface, thus, a

	CSF	GM	WM	All Tissues
T1 only	0.7701 ± 0.0338	0.7483 ± 0.0316	0.8123 ± 0.0325	0.7769 ± 0.0234
T2 FLAIR only	0.7044 ± 0.0081	0.6167 ± 0.0180	0.5865 ± 0.1757	0.6358 ± 0.0531
T1 + T2 FLAIR	**0.7761** ± 0.0248	**0.7539** ± 0.0226	**0.8150** ± 0.0313	**0.7817** ± 0.0233

Table 5: Tissue segmentation results in terms of Dice mean and standard deviation across each patient.

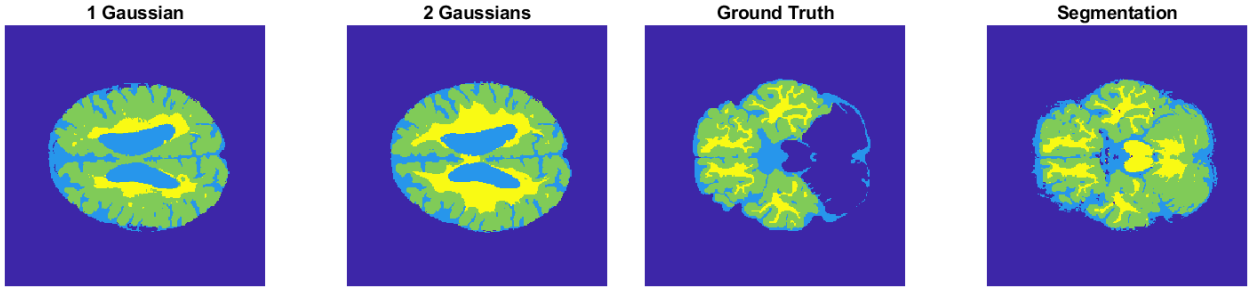

Figure 4: Segmentation results using 1 and 2 Gaussians for WM

Figure 5: Segmentation results for slice 14 of subject 5 (GM shown in green)

faster parameter tuning. The function takes as arguments the path to the MRI volume (T1 or T2 FLAIR or both T1 and T2 FLAIR) and a settings object that stores the parameters to be used for bias field correction. The function returns a struct that holds the file names where the probability maps generated by SPM were stored. The details of the code can be followed by reading the comments in the provided scripts.

Along with this, other methods provided by the SPM library were used for different purposes such as saving the resulting segmentation masks of the tissues in a single NIFTI file.

After obtaining the posterior probability maps, each voxel is assigned to the class with the highest probability. Recall that the tool generates 6 probability maps: CSF, GM, WM, bone, soft tissue and air.

To evaluate the results, the segmentation masks need to be consistent with the provided labels in terms of class indexes. The voxels belonging to skull, soft tissue and air were set to 0 (skull stripping) determining in this way a "background" class. In this way every pixel is classified as belonging to one of four classes: {*Background: 0, CSF: 1, GM: 2, WM: 3*}

Results and discussion

The evaluation of the results was made using the Dice Score (Equation 2). For each subject and for each modality the full volume segmentation was performed. The mean and standard deviation of the metric is computed for each tissue across the five subjects, the results can be seen in table 5. Example results of slice 24 for each patient can be seen in figure 7.

As shown in table 5 and figure 6, a better segmentation can be achieved using the multichannel approach for all tissue types. Even if the median values do dot differ much from using just the T1 sequence, we can appreciate a smaller dispersion of the dice scores obtained. This tell as that providing the algorithm with more information to estimate the class-conditioned voxel intensities distribution (likelihood - GMM) turn into a performance increase.

It can be seen from the results that segmenting T2 FLAIR sequences was more challenging. The table shows a very low result of the mean dice and a larger standard deviation for the WM class using T2. A deeper inspection of the individual results revealed the presence of an outlier when using this modality which is the second subject where the resulting dice was of 0.2736. This can be clearly seen in figure 7 (second row, third column) and in figure 6.

Motivated by this finding, we did a minor parameter tuning for subject 2 and T2-FLAIR. We varied the cleanup parameter from mild to strong cleaning without improvement in the results. Additionally, we varied the number of Gaussians in the set {1,2,3,4}. The Dice Score improved from **0.2736** to **0.5193** when changing the parameter from 1 (default) to 2. Figure 4 shows the segmentation result of slice 24 before and after increasing the number of Gaussians. This shows the potential of parameter tuning, but due to the intensive computations required to find the best set of segmentation parameters for every subject and every

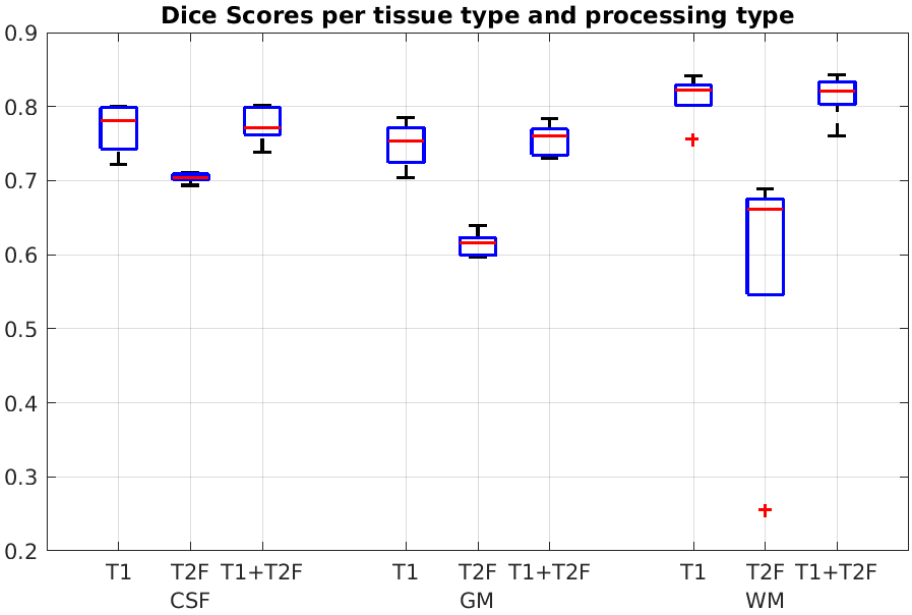

Figure 6: Dice scores per tissue type and processing modality.

modality this approach was not further explored.

In addition to the previous results, we can see in figure 6 that in general the gray matter tissue was more difficult to segment than the other two for all the methods. This could be related to the intrincated anatomical structure it represents, as we can appreciate in figure 5.

Finally it is worth mentioning that more exams should be taken into account to properly validate these results. Running all the experiments across only 5 subjects may have biased the discussed results and did not let us compute the proper statistical significance.

References

[1] J. Ashburner, G. Barnes, C.-C. Chen, J. Daunizeau, G. Flandin, K. Friston, D. Gitelman, V. Glauche, R. Henson, C. Hutton, A. Jafarian, S. Kiebel, J. Kilner, V. Litvak, J. Mattout, R. Moran, W. Penny, C. Phillips, A. Razi, and P. Zeidman, "Spm12 manual," 10 2021.

[2] J. Juntu, J. Sijbers, D. Van Dyck, and J. Gielen, "Bias field correction for mri images," in *Computer Recognition Systems*, M. Kurzyński, E. Puchała, M. Woźniak, and A. żołnierek, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 543–551.

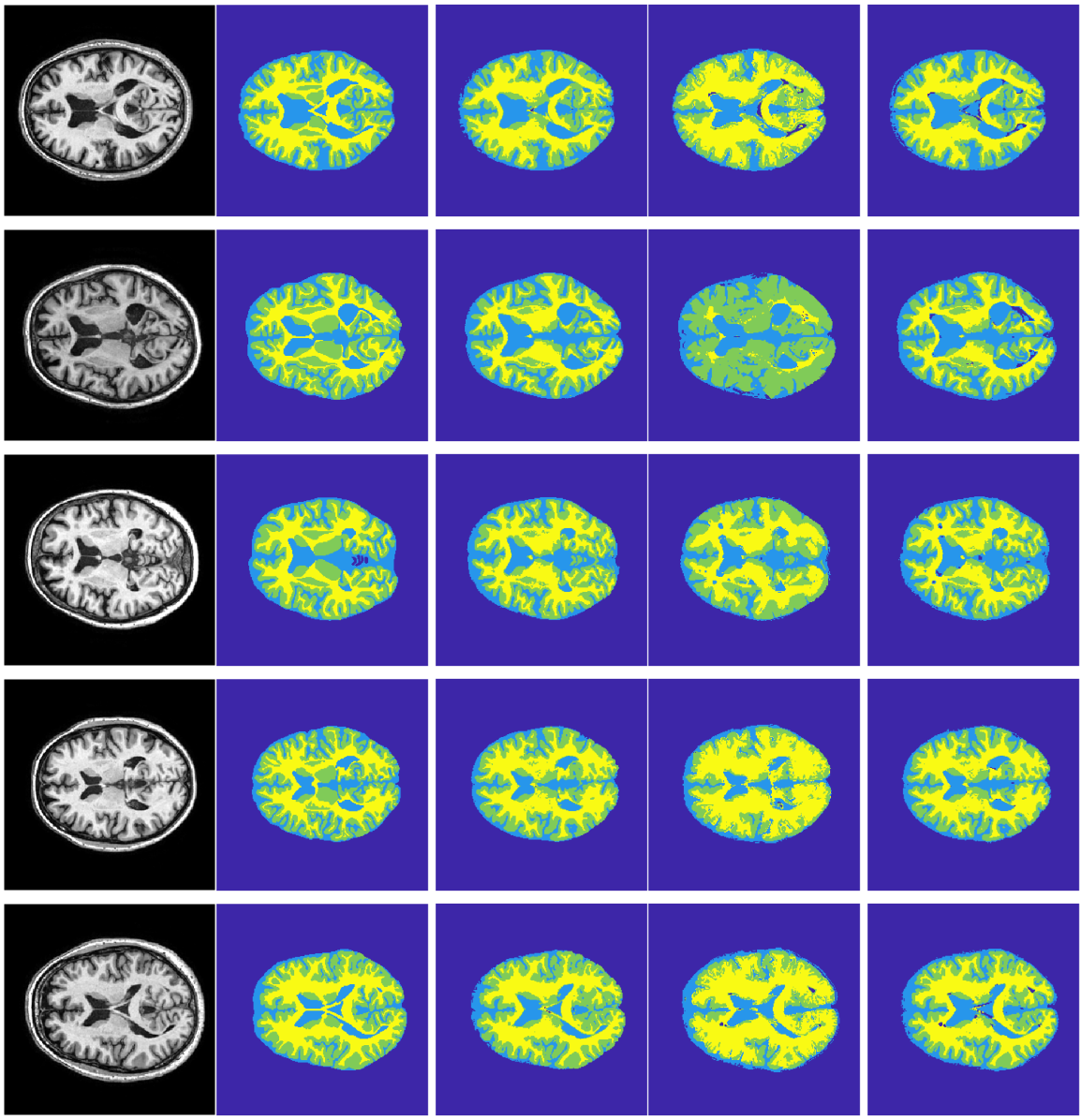

Figure 7: Segmentation results for slice 24, each row depicts a subject from 1 to 5 in order.

Appendix I

I. Bias Field Correction

In the following figures, we show the results of bias field correction with the optimal values for regularisation and FWHM. The figures show the original slice, bias-corrected slice and the bias field for T1 (upper row) and T2 FLAIR (lower row) for each subject.

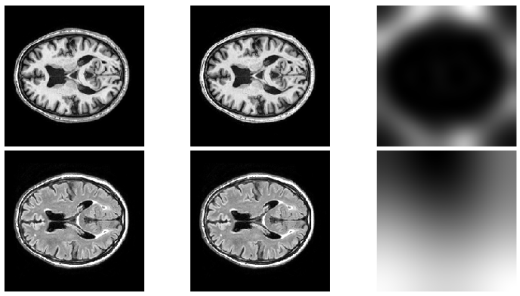

Figure 1: Bias-field correction for subject 1

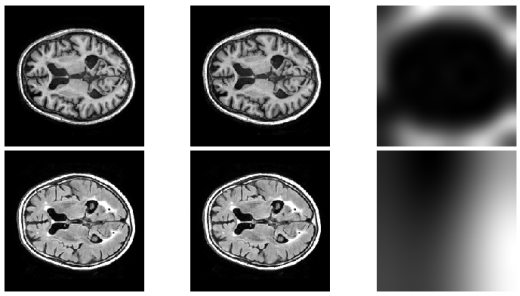

Figure 2: Bias-field correction for subject 2

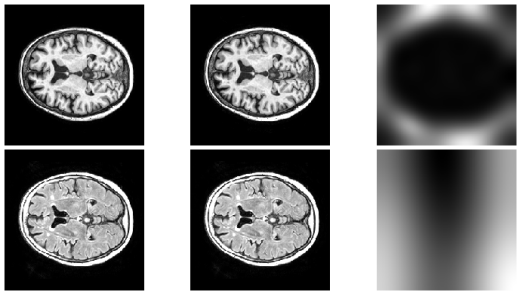

Figure 3: Bias-field correction for subject 3

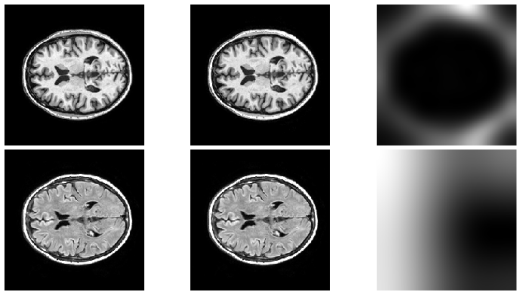

Figure 4: Bias-field correction for subject 4

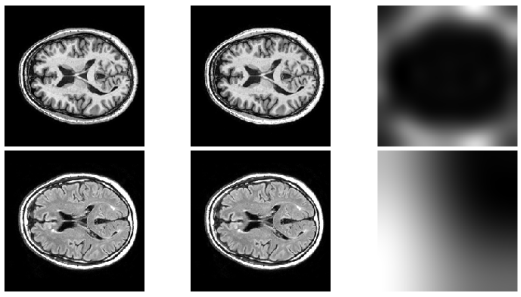

Figure 5: Bias-field correction for subject 5

II. Skull Stripping

In the following figures, we show the results of skull stripping with the optimal values for bias field regularisation and FWHM parameters. The figures show the the results for each subject after an SPM segmentation using T1 only, T2 FLAIR only and T1 with T2 FLAIR respectively.

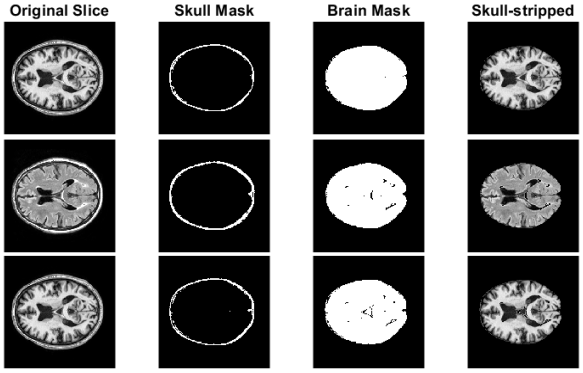

Figure 6: Skull stripping results for subject 1

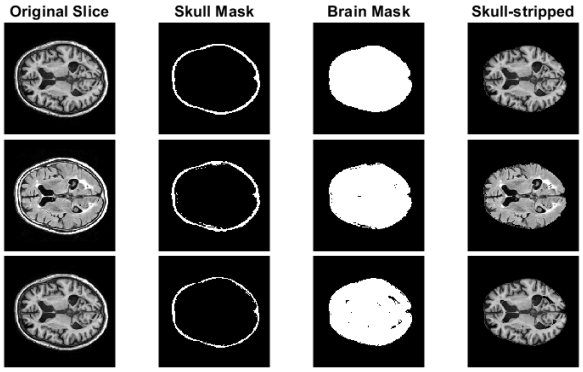

Figure 7: Skull stripping results for subject 2

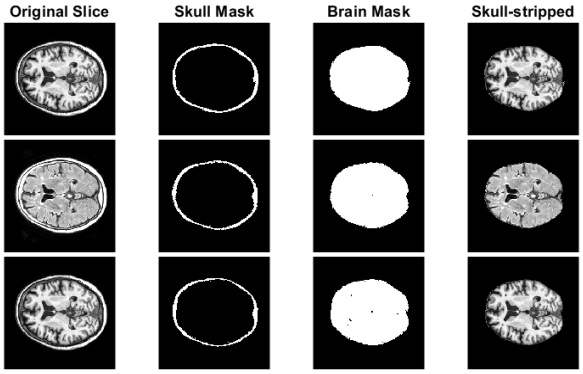

Figure 8: Skull stripping results for subject 3

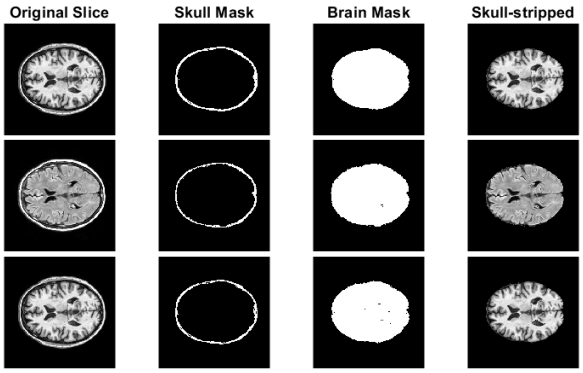

Figure 9: Skull stripping results for subject 4

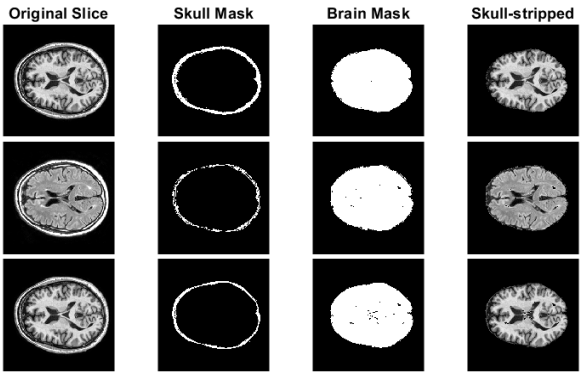

Figure 10: Skull stripping results for subject 5