

Evaluation of an automatic brain segmentation method developed for neonates on adult MR brain images

Pim Moeskops^{a,b}, Max A. Viergever^a, Manon J.N.L. Benders^b, and Ivana Išgum^a

^aImage Sciences Institute, University Medical Center Utrecht, The Netherlands;

^bDepartment of Neonatology, University Medical Center Utrecht, The Netherlands

ABSTRACT

Automatic brain tissue segmentation is of clinical relevance in images acquired at all ages. The literature presents a clear distinction between methods developed for MR images of infants, and methods developed for images of adults. The aim of this work is to evaluate a method developed for neonatal images in the segmentation of adult images. The evaluated method employs supervised voxel classification in subsequent stages, exploiting spatial and intensity information. Evaluation was performed using images available within the MRBrainS13 challenge. The obtained average Dice coefficients were 85.77% for grey matter, 88.66% for white matter, 81.08% for cerebrospinal fluid, 95.65% for cerebrum, and 96.92% for intracranial cavity, currently resulting in the best overall ranking. The possibility of applying the same method to neonatal as well as adult images can be of great value in cross-sectional studies that include a wide age range.

Keywords: Automatic brain segmentation, Supervised voxel classification, MRI

1. INTRODUCTION

Brain tissue segmentation is a prerequisite for many clinical applications in images acquired at all ages, i.e. from preterm born infants to elderly patients. Evaluation frameworks such as NeoBrainS12 (<http://neobrain12.isi.uu.nl>), evaluating neonatal brain segmentation, and MRBrainS13 (<http://mrbrains13.isi.uu.nl>), evaluating adult brain segmentation, show that even though automatic segmentation methods are usually developed for specific age ranges or acquisition protocols, each presenting specific challenges, the underlying techniques share many correspondences. However, none of the participating segmentation methods have been evaluated in both studies.

Popular approaches for brain tissue segmentation are based on atlas registration,^{1,2} pattern recognition,^{3,4} and level sets.^{5,6} Literature describing brain segmentation methods applied to preterm and term-born infants frequently emphasises the differences between segmentation of neonatal and adult images.^{2,4,6} Challenges in neonatal brain segmentation include reduced single-to-noise ratio, reduced contrast between tissue types, increased partial volume effects, and the development of cortical folding and myelination. Because of ongoing myelination, the contrast between white matter and grey matter is inverted in infants. However, the brains of preterm infants already show many similarities with adult brains, and it is therefore likely that very similar methodology can be used to segment both (see Figure 1).

Recently, a method for automatic brain segmentation in preterm infants has been developed, of which a preliminary version was presented by Chiță et al.⁷ The most recent version of this method has been evaluated within the NeoBrainS12 challenge. The aim of the currently presented study is to evaluate whether this method, developed for the segmentation of neonatal MR images, can be employed for the segmentation of adult patients.

Correspondence to Pim Moeskops, E-mail: p.moeskops@umcutrecht.nl, Telephone: +31 88 75 69695

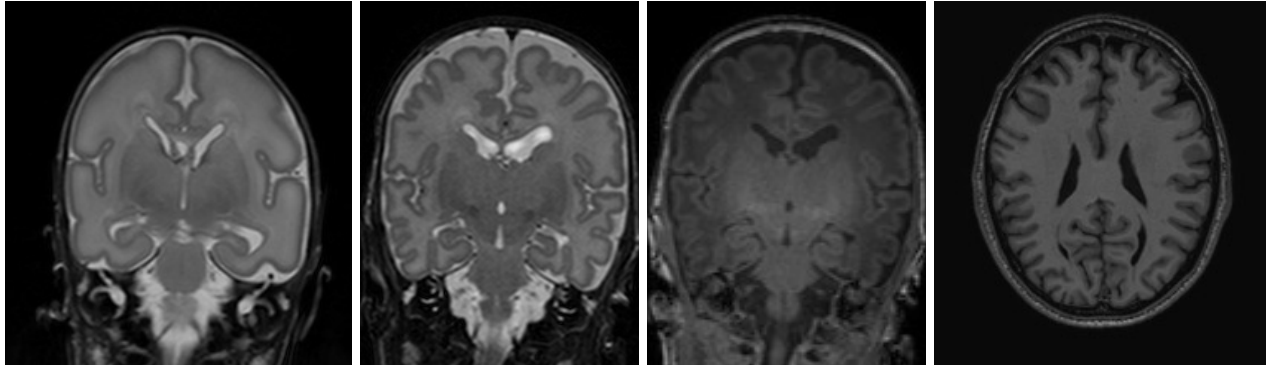


Figure 1. Example MR brain images of preterm born infants and an adult patient. From left to right: a coronal T₂-weighted image of a preterm born infant acquired at 30 weeks postmenstrual age, a coronal T₂-weighted image of a preterm born infant acquired at 40 weeks postmenstrual age, a coronal T₁-weighted image of a preterm born infant acquired at 40 weeks postmenstrual age, and an axial T₁-weighted image of a 67 year old patient.

2. DATA

The images used in this work were obtained from the MRBrainS13 challenge (<http://mrbrains13.isi.uu.nl>). For 20 patients with an average age of 70.5 ± 4.0 years, T₁-weighted images, using turbo field echo (TFE) and inversion recovery (IR) sequences as well as T₂-weighted images using a fluid attenuated (FLAIR) sequence were acquired on a 3T Philips scanner. The images were registered to the FLAIR images, and bias-corrected using SPM8.⁸

Manual annotations of cortical grey matter, basal ganglia, white matter, white matter lesions, cerebrospinal fluid in the extracerebral space, ventricles, cerebellum, and brainstem were performed in these images using a freehand spline drawing tool. The challenge evaluates three combined tissue classes: cortical grey matter and basal ganglia (GM), white matter and white matter lesions (WM), and cerebrospinal fluid in the extracerebral space and ventricles (CSF). Additionally, segmentations of cerebrum and intracranial cavity are evaluated. Five images with manual segmentations are provided as training data, the other fifteen are used to test the performance of the participating algorithms.

Because the evaluated method performed neonatal segmentation based on T₂-weighted images only, for the purpose of adult segmentation only the TFE T₁-weighted images (Table 1) were used, because they provide the most similar contrast to neonatal T₂-weighted images.

Table 1. Acquisition parameters for the T₁-weighted images. Voxel sizes and dimensions are presented in, respectively, right-left, anterior-posterior, and inferior-superior directions.

Sequence	Turbo Field Echo
Acquisition protocol	3D
Repetition time [ms]	7.9
Echo time [ms]	4.5
Scan time [min]	6.73
Field of view [mm ³]	192 × 232 × 256
Acquisition resolution	192 × 256 × 256
Acquired voxel size [mm ³]	1.0 × 1.0 × 1.0
Registered resolution	240 × 240 × 48
Registered voxel size [mm ³]	0.96 × 0.96 × 3.0

3. METHODS

The method performed the segmentation of GM, WM and CSF in three subsequent segmentation stages. The training data provided by the MRBrainS13 challenge (five images) were used to train the method, following the requirements of the challenge.

3.1 Brain mask

Prior to the voxel classification, the region of interest was selected by generating a brain mask. These brain masks were obtained using multi-atlas registration of the training images. All five training images were registered to each test image, in an affine and an elastic step using elastix.⁹ Both registrations were performed using a multi-resolution approach (coarse to fine) with mutual information as metric and a gradient descent optimiser. The brain mask was obtained by majority voting, i.e. at least three of the five registered brain masks should have labelled the voxel as part of the brain. Only voxels within this mask were further analysed.

3.2 First segmentation stage

For all voxels within the brain mask, features based on spatial characteristics and intensity of the T₁-weighted image were computed. The spatial features were described as normalised coordinates relative to the brain mask and the shortest Euclidean distance to boundary of the mask. The intensity characteristics were computed by applying Gaussian filters up to and including the second order to the T₁-weighted image, using scales $\sigma = 0.5, 1.0, 1.5,$ and 2.0 mm. Prior to the classification, all features were normalised to zero mean and unit variance. Voxel classification was performed using three independent two-class k -nearest neighbour (k NN) classifiers with 101 neighbours, one for each tissue type. Feature selection for these classifiers was performed based on neonatal data; the same features were used in this study.

Based on the posterior probability of each classifier, the voxels were either assigned to one of the tissues (WM, GM, or CSF), or were selected for the reclassification in the second segmentation stage. This means that the voxels with a high probability for one of the tissue types were labelled as such, and the voxels with a low probability for all tissue types were labelled as background. All other voxels were reanalysed in the second stage. The upper and lower thresholds were selected based on experiments with neonatal data.

3.3 Second segmentation stage

The second segmentation stage used, in addition to the previously described spatial and intensity features, features based on the output of the first stage. To this end, Gaussian filters up to and including the second order, using scales of $\sigma = 0.5$ and 1.0 mm, were applied to the posterior probabilities of the first stage classifiers. All features were normalised to zero mean and unit variance. Based on this combined feature set, three independent two-class support vector machine (SVM) classifiers were trained.

To obtain a binary segmentation, the probabilistic result for each tissue type was thresholded at 0.5. Because voxel classification could result in spatially isolated (clusters of) voxels, only the largest connected component (in 3D) was retained for the WM as well as the GM segmentation. The CSF segmentation included both the ventricular CSF and the extracerebral CSF, therefore none of the voxels labelled as CSF were discarded.

Because dedicated classification was performed for each tissue type separately, some voxels could be labelled as more than one tissue type, or remain unlabelled. These voxels, typically located along tissue boundaries, were corrected in the third segmentation stage.

3.4 Third segmentation stage

The third segmentation stage used the same features as the second stage, using updated probabilistic results resulting from the second stage. Additionally, the shortest Euclidean distances to the each of the three binary segmentations, resulting from the second stage, were added as features. All features were normalised to zero mean and unit variance. These features were input for a four-class (WM, GM, CSF and background) k NN classifier, using 55 neighbours. Again, the same set of features was used as for the neonatal data.

All voxels reanalysed in the third stage were labelled to the tissue type with the highest posterior probability and combined with the binary result of the second stage.

4. RESULTS

The evaluation results according to the MRBrainS13 evaluation framework are listed in Table 2, and an example automatic segmentation is shown in Figure 2. The method achieved the highest Dice coefficient and the lowest modified Hausdorff distance for GM, WM, and cerebrum, resulting in the rank of overall best performing method.

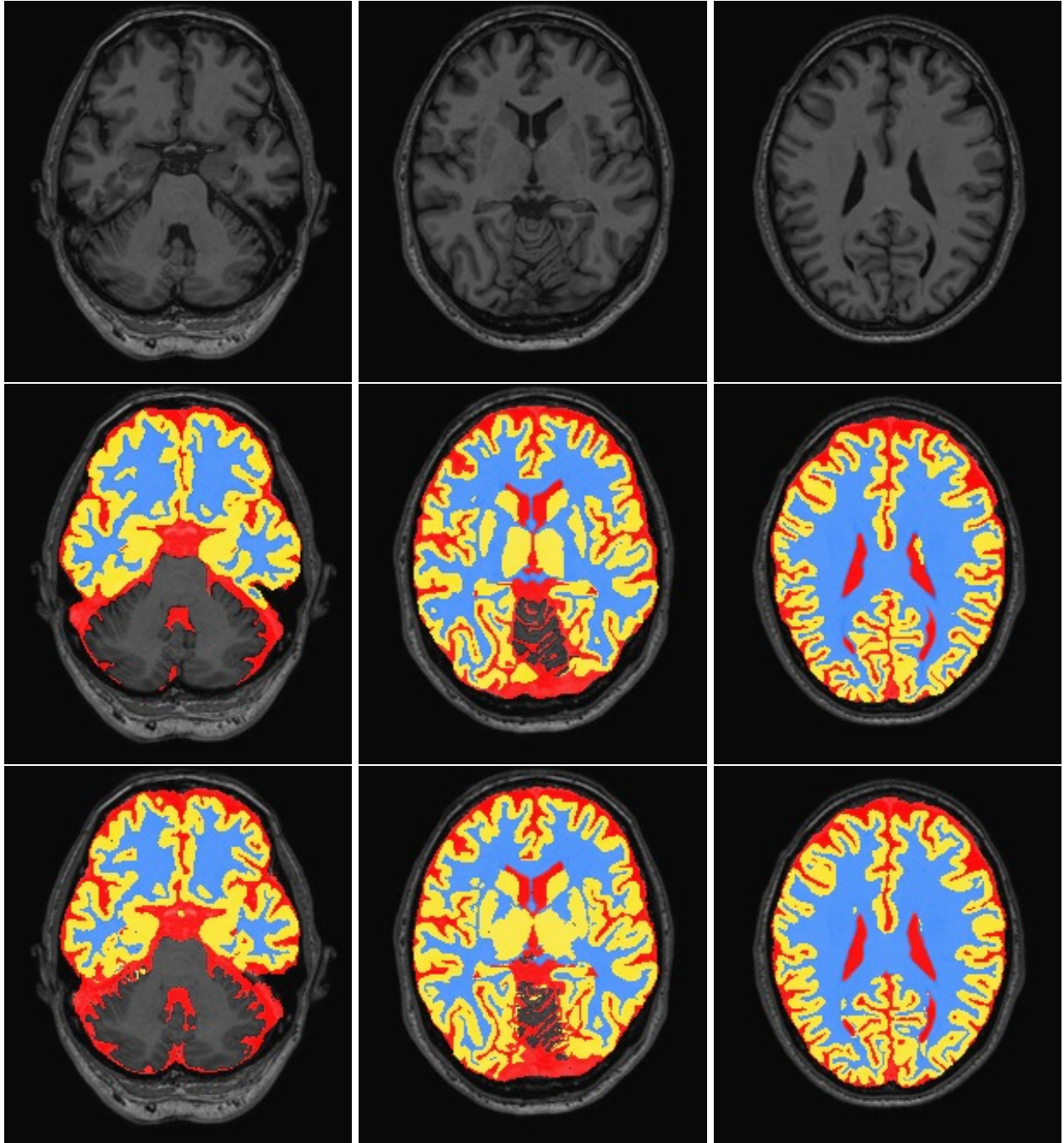


Figure 2. Example segmentation result for one slice. From top to bottom: the T_1 -weighted image, the reference segmentation, and the automatic segmentation. GM is shown in yellow, WM is shown in blue, and CSF is shown in red.

Table 2. Average results obtained within the MRBrainS13 evaluation framework for the presented method and the best ranked method for each tissue type in terms of Dice coefficient, modified Hausdorff distance (MHD), and absolute volume difference (AVD). The cerebrum was evaluated as the GM and WM segmentations combined, and the intracranial cavity was evaluated as the GM, WM, and CSF segmentations combined.

	Presented Method			Best method per tissue type (MRBrainS13)		
	Dice [%]	MHD [mm]	AVD [%]	Dice [%]	MHD [mm]	AVD [%]
Grey matter	85.77	1.62	6.62	84.12	1.92	5.44
White matter	88.66	2.06	6.96	88.42	2.36	6.02
Cerebrospinal fluid	81.08	2.66	9.77	82.18	2.46	8.81
Cerebrum	95.65	2.18	2.80	95.12	2.74	3.24
Intracranial cavity	96.92	4.29	1.74	97.92	2.55	0.91

5. DISCUSSION

This paper presented evaluation of an automatic brain segmentation method developed for preterm infants on images of adult patients. Even though previous work emphasised difficult application of methods developed for adult images to segmentation of the neonatal brain, we have shown that a method developed for neonatal brain segmentation can be successfully applied to the segmentation of adult brain images. This may be due to the challenging segmentation of (preterm) infants, requiring analysis focusing on subtle changes in the image intensities in the brain.

In the presented segmentation experiments, the parameter settings were identical to the parameters defined with neonatal data. Namely, the same features, classifiers and their settings as well as thresholds on the posterior probabilities were used. However, the features were extracted from T_1 - instead of T_2 -weighted images, because of the inverted contrast of WM and GM between neonatal and adult images. This suggests that the presented approach is robust to changes in the acquisition protocol, as well as to the anatomical differences between preterm infants and adults. Nevertheless, dedicated parameter tuning and feature selection, might possibly further improve the result for these images.

A difference with the neonatal segmentation approach was the brain mask, which was obtained using multi-atlas registration instead of FSL's brain extraction tool (BET), which produced insufficiently consistent brain masks for these data. An advantage of BET, opposed to multi-atlas registration, is that it does not require training data.

Even though the MRBrainS13 challenge provided T_1 -weighted IR and T_2 -weighted FLAIR acquisitions as well, the automatic segmentation was performed using standard T_1 -weighted images only. Performing segmentation using only one image allows omitting registration between the images, and therefore eliminates possible registration errors.

Following the requirements of the MRBrains13 challenge, the method has been tested for the labelling of GM, WM and CSF. Including other tissue types in the segmentation is an interesting topic for further research. Ventricular and extracerebral CSF, as well as cortical and deep GM, i.e. the basal ganglia, were combined in single tissue types. For the neonatal data, ventricular CSF and deep GM were not included in the CSF and GM segmentations, but labelled as background. This suggests that even though the evaluated method was not specifically designed to perform these tasks, it is able to adapt based on new training data describing these different classes.

6. CONCLUSION

The evaluated method, developed for the segmentation of neonatal MR brain images, obtained accurate segmentation and achieved the highest Dice coefficient and the smallest modified Hausdorff distance for GM, WM and cerebrum in the MRBrainS13 challenge on adult brain segmentation. The possibility of using the same segmentation method from very early neonatal until old adult age can be of great value in performing cross-sectional studies including patients of a wide age range and possibly long-term longitudinal analysis.

REFERENCES

- [1] Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M., “Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.,” *Neuron* **33**(3), 341–355 (2002).
- [2] Makropoulos, A., Gousias, I., Ledig, C., Aljabar, P., Serag, A., Hajnal, J., Edwards, A. D., Counsell, S., and Rueckert, D., “Automatic whole brain MRI segmentation of the developing neonatal brain.,” *IEEE Trans Med Imaging* **33**(9), 1818–1831 (2014).
- [3] Vrooman, H. A., Cocosco, C. A., van der Lijn, F., Stokking, R., Ikram, M. A., Vernooij, M. W., Breteler, M. M. B., and Niessen, W. J., “Multi-spectral brain tissue segmentation using automatically trained k-nearest-neighbor classification.,” *Neuroimage* **37**(1), 71–81 (2007).
- [4] Anbeek, P., Išgum, I., van Kooij, B. J., Mol, C. P., Kersbergen, K. J., Groenendaal, F., Viergever, M. A., de Vries, L. S., and Benders, M. J., “Automatic segmentation of eight tissue classes in neonatal brain MRI,” *PloS ONE* **8**(12), e81895 (2013).
- [5] Baillard, C., Hellier, P., and Barillot, C., “Segmentation of brain 3D MR images using level sets and dense registration.,” *Med Image Anal* **5**(3), 185–194 (2001).
- [6] Wang, L., Shi, F., Li, G., Gao, Y., Lin, W., Gilmore, J. H., and Shen, D., “Segmentation of neonatal brain MR images using patch-driven level sets.,” *Neuroimage* **84**, 141–158 (2014).
- [7] Chiță, S., Benders, M. J., Moeskops, P., Kersbergen, K. J., Viergever, M. A., and Išgum, I., “Automatic segmentation of the preterm neonatal brain with MRI using supervised classification,” in [*SPIE Medical Imaging*], 86693X–86693X, International Society for Optics and Photonics (2013).
- [8] Friston, K. J., Ashburner, K., Kiebel, K., Nichols, T., and Penny, W., [*Statistical Parametric Mapping: The Analysis of Functional Brain Images*], Academic Press (2007).
- [9] Klein, S., Staring, M., Murphy, K., Viergever, M. A., and Pluim, J. P. W., “elastix: a toolbox for intensity-based medical image registration.,” *IEEE Trans Med Imaging* **29**(1), 196–205 (2010).