Automated Brain-Tissue Segmentation by Multi-Feature SVM Classification

Annegreet van Opbroek[1], Fedde van der Lijn[1] and Marleen de Bruijne[1,2]

[1] Biomedical Imaging Group Rotterdam, Departments of Medical Informatics and Radiology, Erasmus MC - University Medical Center Rotterdam, the Netherlands
[2] Department of Computer Science, University of Copenhagen, Denmark

Abstract. We present a method for automated brain-tissue segmentation through voxelwise classification. Our algorithm uses manually labeled training images to train a support vector machine (SVM) classifier, which is then used for the segmentation of target images. The classification incorporates voxel intensities from a T1-weighted scan, an IR scan, and a FLAIR scan; features to encode the voxel position in the image; and Gaussian-scale-space features and Gaussian-derivative features at multiple scales to facilitate a smooth segmentation.
An experiment on data from the MRBrainS13 brain-tissue-segmentation challenge showed that our algorithm produces reasonable segmentations in a reasonable amount of time.

1 Introduction

The segmentation of brain images can provide useful information about neurogenerative diseases such as dementia and multiple sclerosis, which is useful both in research and clinical practice. Manual segmentation of brain images is a tedious task however, which is why a variety of methods have been developed for automated segmentation.

Three types of automated brain-tissue-segmentation techniques can be distinguished: techniques that use manually segmented images to train a segmentation algorithm, techniques that require no manually segmented training images, and techniques that use atlases. Methods that fall in the first category are usually based on supervised classification. In supervised classification labeled training data is used to extract features and train a classifier, such as a k-nearest neighbor (kNN) classifier [1], a random decision forest [2], or a support vector machine (SVM) [3]. The labeled training data used in the classification is usually obtained with the same scanner as the target data, but Van Opbroek et al. [3] propose a transfer-learning SVM that can deal with labeled training data from other scanners.

Methods that do not require manually labeled training data include clustering algorithms such as the fuzzy c-means algorithm [4] and mean-shift clustering [5].

A third option is to obtain a segmentation with the help of manually constructed atlases, which can be used to automatically select and label training

data from a target image [6], to initialize an expectation-maximization algorithm [7], or to improve parameter estimation in classification with an EM algorithm [8].

In this paper we present a brain-tissue-segmentation framework based on supervised voxelwise classification with an SVM classifier, which is a state-of-the-art machine-learning classifier. In contrast to other techniques [1][2][4][5] our segmentation scheme uses Gaussian-scale-space features and Gaussian-derivative features next to the image intensities, to facilitate a smooth segmentation of the different tissues. The performance of our method has been tested on 12 images from the brain-tissue segmentation challenge MRBrainS13[3].

2 Material and Methods

2.1 MRBrainsS13 Training and Test Data

Training and test images have been acquired at the UMC Utrecht in the Netherlands and concern patients with diabetes and matched controls (age>50) with varying degrees of atrophy and white-matter lesions. From all subjects a T1-weighted scan, a T1-weighted inversion recovery (IR) scan, and a fluid attenuated inversion recovery (FLAIR) scan have been obtained, all with $0.958 \times 0.958 \times 3.0$ mm^3 voxel size. The three sequences have been registered and the images have been bias corrected.

Five images were provided to train a segmentation algorithm. The training images were manually segmented into background and eight tissues: cortical gray matter (GM), basal ganglia, white matter (WM), white-matter lesions, cerebrospinal fluid (CSF) in the extracerebral space, the ventricles, the cerebellum, and the brainstem. Also, twelve test images were provided which had to be segmented into background, gray matter (cortical gray matter + basal ganglia), white matter (white matter + white-matter lesions), and cerebrospinal fluid (CSF in the extracerebral space + ventricles). Segmentation of the cerebellum and the brainstem was not included in the evaluation.

2.2 Preprocessing

All images were normalized by a range-matching procedure that scaled the voxels within the mask so that the 4th percentile was mapped to zero, and the 96th percentage to one. Since the images of the challenge were already bias corrected, no further bias correction was performed.

2.3 Brain Segmentation

For the test images brain masks were created with multi-atlas segmentation. As atlases we used the five T1-weighted training scans, both in the original and in a left-right-flipped version. Brain masks were obtained for these ten atlases

[3] http://mrbrains13.isi.uu.nl

by binarizing the training labels (including the brainstem and the cerebellum). Each atlas image was registered to the unlabeled test images using Niftyreg [9] by computing an affine transformation, followed by a non-rigid deformation using a 5mm B-spline grid and normalized mutual information. A final mask was computed using STEPS [10]. This method deforms both atlas images and labels, selects per voxel location the five most similar atlases (based on local normalized cross correlation), and fuses their labels using STAPLE [11].

2.4 Features

From each of the three sequences, T1, IR, and FLAIR, 10 features were extracted:

- The intensity
- The intensity after convolution with a Gaussian kernel with $\sigma = 1, 2, 3$ mm^3
- The gradient magnitude of the intensity after convolution with a Gaussian kernel with $\sigma = 1, 2, 3$ mm^3
- The laplacian of the intensity after convolution with a Gaussian kernel with $\sigma = 1, 2, 3$ mm^3.

The spatial information was incorporated by adding three spatial features: the x, y, and z coordinate of a voxel, divided by respectively the length, width, and height of the brain. This resulted in a total of 33 features. All features were scaled to zero mean and unit standard deviation.

2.5 SVM Classification

We performed supervised voxelwise classification with a soft-margin SVM [12]. An SVM learns a decision function $f(\boldsymbol{x}) = \boldsymbol{v} \cdot \phi(\boldsymbol{x}) + v_0$, where the model parameters $\boldsymbol{v}$ and v_0 are determined from the training data, and ϕ is a mapping to project a sample $\boldsymbol{x}_i$ into a feature space $\phi(\boldsymbol{x}_i)$. This mapping defines a kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle$ by means of the kernel trick.

The model parameters $\boldsymbol{v}$ and v_0 are determined from the training data by optimizing the following criterion

$$\begin{aligned} \min_{\boldsymbol{v}} \ & \frac{1}{2}\|\boldsymbol{v}\|^2 + C\sum_{i=1}^{N} \xi_i \\ \text{s.t. } & y_i \boldsymbol{v}^T \phi(\boldsymbol{x}_i) + v_0 \geq 1 - \xi_i \\ & \xi_i \geq 0 \ . \end{aligned} \tag{1}$$

The first term, $\|\boldsymbol{v}\|^2$ maximizes the margin around the decision function, and $C\sum_{i=1}^{N} \xi_i$ minimizes the number of training samples that are either misclassified or lie within the margin. C is a parameter to trade off between maximizing the margin and minimizing $\sum_i \xi_i$ where a sample receives a value $\xi_i > 1$ if it is misclassified, a value $0 < \xi_i \leq 1$ if it is correctly classified but lies within the margin, and a value $\xi_i = 0$ otherwise.

We performed six-class classification to classify cortical GM, basal ganglia, WM, white-matter lesions, CSF in the extracerebral space, and the ventricles. Since SVMs are designed for two-class classification, the classification was extended to multi-class classification by one-versus-one classification which means that a total of 15 SVMs were trained to distinguish between the six classes.

For the SVM classification an implementation in LIBSVM [13] was used.

Classification Parameters To enable for non-linear decision functions a radial basis kernel was chosen. A suitable value for the SVM parameter C and the kernel parameter γ were determined in a grid-search experiment on the five training images in which the total classification error was minimized. This resulted in values of $C = 8$ and $\gamma = 0.01$.

The SVM classifier was trained on a total of 10 000 samples per training image, which were randomly selected from inside the provided brain mask excluding the cerebellum and the brain stem.

2.6 Postprocessing

Postprocessing involved fusing the voxels segmented as cortical gray matter and basal ganglia, white matter and white-matter lesions, and CSF in the extracerebral space and the ventricles. Subsequently a closing algorithm was performed on the regions segmented as CSF, by a dilation of 1 voxel, followed by an erosion of 1 voxel. This was done to reduce the amount of voxels in the CSF that were segmented as WM or GM.

2.7 Outcome Measures

Segmentation results on the 12 test images were compared to the manual segmentations based on three measures:

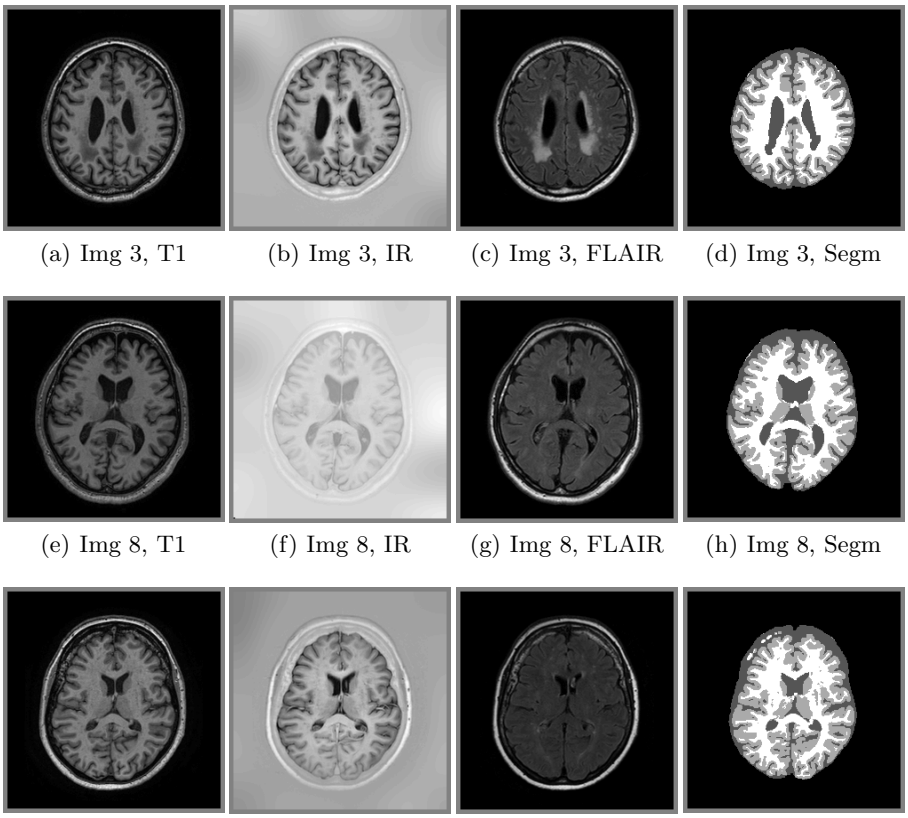
- The DICE overlap [14]
- The modified Hausdorff distance (MHD) [15]
- The absolute volume difference (AVD) [16].

3 Results

Table 3 shows the mean and standard deviation (std) of the scores. The evaluation measures were calculated for 5 tissues: GM, WM, CSF, brain (WM+GM), and all intracranial structures (WM+GM+CSF). For the DICE score our algorithm scored best on white-matter segmentation, with a mean DICE coefficient of 88.3%, and slightly lower for gray matter, with a mean DICE of 84.5%. The most errors were made in the CSF, which had a mean DICE score of only 78.0%. Also on the other two scores, MHD and AVD, white matter and gray matter had a better score than CSF.

Structure	DICE (%) Mean	Std	MHD (mm) Mean	Std	AVD (%) Mean	Std
Gray Matter	84.51	1.44	1.97	0.34	6.92	3.09
White Matter	88.30	0.98	2.41	0.50	6.79	5.19
Cerebrospinal fluid	78.00	5.43	3.31	0.80	25.98	18.49
Brain	95.05	0.53	2.79	0.82	3.51	1.61
All Intracranial Structures	95.86	1.32	4.02	1.20	5.67	3.14

Table 1. Results on 12 test images: mean and standard deviation (std) for GM, WM, CSF, brain (WM+GM), and all intracranial structures (WM+GM+CSF).



(a) Img 3, T1 (b) Img 3, IR (c) Img 3, FLAIR (d) Img 3, Segm

(e) Img 8, T1 (f) Img 8, IR (g) Img 8, FLAIR (h) Img 8, Segm

(i) Img 10, T1 (j) Img 10, IR (k) Img 10, FLAIR (l) Img 10, Segm

Fig. 1. T1, IR, FLAIR images and segmentations (Segm) of 3 of the 12 test images. Image 3 in (a)-(d) contains a large amount of lesions, image 8 in (e)-(h) was given the overall best score, image 10 in (i)-(l) was given the overall worst score.

Example segmentations for slices from images 3, 8, and 10 are presented in Fig. 1. Fig. 1(a),(e),(i) show the T1-weighted images, Fig. 1(b),(f),(j) show the T1-weighted IR scans, Fig. 1(c),(g),(k) show the FLAIR scans, and Fig. 1(d),(h),(l) show the segmentations. Image 3 in Fig. 1(a)-(d) has a very large amount of lesions, Image 8 in fig. 1(e)-(h) is the image that overall scored the best, and image 10 in Fig. 1(i)-(l) overall scored worst.

For all images the determined brain mask was too large in the front, as can be seen in the segmentations in Fig. 1(d),1(h),(l). In most images this led to voxels in the front of the image being erroneously classified as either white matter or gray matter tissue. This effect is most prominent in Fig. 1(l), where WM and GM clusters have appeared in the CSF, but it can also been seen in the segmentations in Fig. 1(d),1(h). In images with a large amount of lesions, as in Image 3, lesion voxels were sometimes erroneously classified as CSF, as is shown in Fig. 1(d).

4 Discussion

We have presented an algorithm for automated brain extraction and brain-tissue segmentation. The brain-extraction algorithm is based on multi-atlas segmentation with the STEPS [10] algorithm; the tissue classification is based on voxelwise SVM classification. In the voxelwise classification T1, IR, and FLAIR intensities, spatial features, Gaussian-scale-space features and Gaussian-derivative features were used.

Our algorithm produced reasonable segmentations which were generally quite smooth without further spatial regularization because of the use of the Gaussian-scale-space features and the Gaussian-derivative features. In some slices however, mainly around the basal ganglia, the segmentations were not completely smooth, which was caused by the low contrast between the basal ganglia and the surrounding white matter.

The largest errors were made in the segmentation of the CSF in the extracerebral space, which was mainly because the determined brain mask was too big in the frontal lobe due to a slight misregistration of the atlases. As a result, skull voxels were incorporated in the brain mask, which were sometimes segmented as white or gray matter. As a post-processing step a closing algorithm was performed on the CSF tissue, which segmented some of these voxels as CSF. We believe that refining the masks by including a background class in the SVM classification may improve the results.

Other weaknesses of our algorithm are a slight under segmentation of the basal ganglia, and the misclassification of voxels in the center of large white-matter lesions that are close to the ventricles, which were erroneously segmented as CSF. This second problem could be resolved by excluding the lesions from the tissue segmentation and including a separate lesion-classification step. This separate classification step allows for the exclusion of spatial features from the feature set, which can be very misleading features for lesion segmentation when only a small number of training images is available.

The total runtime of our algorithm per test image was 10 times 8 minutes to perform the registrations for the image mask, 25 seconds to determine the image features, 1.5 minutes to train the SVM classifiers (note that this only needs to be done once for segmentation of all images), and 35 minutes for classification of the test image. The registrations for the image mask were computed on a cluster with AMD Opteron 2216 2.4GHz nodes without multi-threading, the rest was implemented in Matlab and computed on a machine with an Intel Xeon E5620 2.40 GHz CPU. For a voxelwise classification 35 minutes is relatively long, which is due to the fact that a total of 15 one-vs-one SVMs were calculated. The calculation time could be drastically reduced by training on only three labels: gray matter, white matter, and CSF. In a cross-validation experiment on the training set this only slightly increased the mean classification error from 13.9% to 14.3%, but decreased the calculation time with approximately a factor of 5.

To conclude, the proposed multi-feature SVM classification produces reasonable segmentations in a reasonable amount of time. We believe that if the registrations of the training masks to the target images could be improved, and a separate lesion-segmentation algorithm could be included in the segmentation, the resulting segmentations would come close to those of human observers.

5 Acknowledgments

This research is financed by The Netherlands Organization for Scientific Research (NWO).

References

1. Anbeek, P., Vincken, K., Van Bochove, G., Van Osch, M., van der Grond, J., et al.: Probabilistic segmentation of brain tissue in MR imaging. Neuroimage **27**(4) (2005) 795
2. Yi, Z., Criminisi, A., Shotton, J., Blake, A.: Discriminative, semantic segmentation of brain tissue in MR images. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009 (2009) 558–565
3. van Opbroek, A., Ikram, M., Vernooij, M., de Bruijne, M.: Supervised image segmentation across scanner protocols: A transfer learning approach. Machine Learning in Medical Imaging (2012) 160–167
4. Bazin, P., Pham, D.: Topology-preserving tissue classification of magnetic resonance brain images. Medical Imaging, IEEE Transactions on **26**(4) (2007) 487–496
5. Mayer, A., Greenspan, H.: An adaptive mean-shift framework for MRI brain segmentation. Medical Imaging, IEEE Transactions on **28**(8) (2009) 1238–1250
6. Cocosco, C., Zijdenbos, A., Evans, A.: A fully automatic and robust brain MRI tissue classification method. Medical Image Analysis **7**(4) (2003) 513–527
7. Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: Automated model-based tissue classification of MR images of the brain. Medical Imaging, IEEE Transactions on **18**(10) (1999) 897–908
8. Ashburner, J., Friston, K.: Unified segmentation. Neuroimage **26**(3) (2005) 839–851

9. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer methods and programs in biomedicine **98**(3) (2010) 278–284
10. Cardoso, M.J., Leung, K., Modat, M., Keihaninejad, S., Cash, D., Barnes, J., Fox, N.C., Ourselin, S.: Steps: Similarity and truth estimation for propagated segmentations and its application to hippocampal segmentation and brain parcelation. Medical Image Analysis (2013)
11. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. Medical Imaging, IEEE Transactions on **23**(7) (2004) 903–921
12. Cortes, C., Vapnik, V.: Support-vector networks. Machine learning **20**(3) (1995) 273–297
13. Chang, C., Lin, C.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) **2**(3) (2011) 27
14. Dice, L.: Measures of the amount of ecologic association between species. Ecology **26**(3) (1945) 297–302
15. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing images using the hausdorff distance. Pattern Analysis and Machine Intelligence, IEEE Transactions on **15**(9) (1993) 850–863
16. Babalola, K.O., Patenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., Smith, S., Cootes, T., Jenkinson, M., Rueckert, D.: An evaluation of four automatic methods of segmenting the subcortical structures in the brain. Neuroimage **47**(4) (2009) 1435–1447