MS Lesion Segmentation using a Multi-channel Patch-based Approach with Spatial Consistency

Roey Mechrez[1], Jacob Goldberger[2] and Hayit Greenspan[1]

[1]Department of Biomedical Engineering, Tel-Aviv University, Israel
[2]Faculty of Engineering, Bar Ilan University, Ramat-Gan, Israel

ABSTRACT

This paper presents an automatic method for segmentation of Multiple Sclerosis (MS) in Magnetic Resonance Images (MRI) of the brain. The approach is based on similarities between multi-channel patches (T1, T2 and FLAIR). An MS lesion patch database is built using training images for which the label maps are known. For each patch in the testing image, k similar patches are retrieved from the database. The matching labels for these k patches are then combined to produce an initial segmentation map for the test case. Finally a novel iterative patch-based label refinement process based on the initial segmentation map is performed to ensure spatial consistency of the detected lesions. A leave-one-out evaluation is done for each testing image in the MS lesion segmentation challenge of MICCAI 2008. Results are shown to compete with the state-of-the-art methods on the MICCAI 2008 challenge.

1. INTRODUCTION

Patch-based methods for segmentation have been shown to be an effective approach for labeling brain structures (and other body structures), as shown for example in,[1,2] and recently for lesions in.[3] In general, these approaches label each voxel of a target image by comparing the image patch centered on the voxel with patches from an atlas library, and assign the most probable label according to the closest matches. Often, a localized search window centered around the target voxel, is used. Various patch-based label fusion procedures have been proposed and were shown to produce accurate and robust segmentation. Using affine registration, comparable results are shown to works that use non-rigid registration. The focus of the segmentation task in existing works is on regions of interest which are substantial in size and/or are anatomically localized. Thus, a good approximate alignment of the image to an atlas is needed and is a key component. In the current work we extend the investigation of patch-based approaches to the task of MS lesion detection and segmentation in MR images of the brain. The challenges we face include the following: MS lesions are very small. They can appear anywhere in the brain image, and no atlas currently exists for MS lesion data. Examples of MS lesions are presented in Fig. 3.

MS is the most common non-traumatic neurological disease in young adults. It is an inflammatory demyelinating disease that is primarily associated with axonal loss and formation of lesions in the central nervous system, which are characterized by demyelination, axonal injury and axonal conduction block. These classically-described white matter (WM) lesions are visible in conventional magnetic resonance imaging, appearing hyperintense in T2-weighted (T2w) images and hypointense in T1-weighted (T1w) images. Fluid attenuated inversion recovery (FLAIR) images have been shown to be the most sensitive to WM lesions, but can also present other hyperintensity artifacts.[4]

Manual segmentation of WM lesions is a time-consuming process. Furthermore, the 3D data of an MRI scan requires multi-slice segmentation which makes the manual segmentation an overall untangeable task.[5] Fully automated algorithms for MS lesion detection and segmentation have been the focus of research for many years. Due to the complexity of the task and the recent understanding that the task is yet unsolved, the performance of these algorithms is still being challenged and new algorithmic approaches have recently been introduced.[3,6–9] Recent methods have been developed through the MICCAI08 grand challenge.[10] This challenge is helping to advance the field by providing a dataset for evaluation and comparison across the developed techniques.

In the current work we introduce a novel patch-based approach for MS lesion detection and segmentation. The integration of the spatial information to this method is the main contribution of the work. The proposed framework segments the MS image without requiring registration and with no need for an atlas.

Medical Imaging 2015: Image Processing, edited by Sébastien Ourselin, Martin A. Styner, Proc. of SPIE Vol. 9413, 94130O · © 2015 SPIE · CCC code: 1605-7422/15/$18 · doi: 10.1117/12.2082558

2. PATCH-BASE SEGMENTATION AND LABEL REFINEMENT

Our method is based on labeling the test image voxels (as lesion or non-lesion) by finding similar patches in a database of manually labeled images. The training step involves constructing a patch database using expert-marked lesion regions which provides voxel-level labeling. Given a test image, we first detect candidate lesion regions for further analysis. For all candidate voxels we find the k-nearest neighbor patches from the patch database. Labels of the selected patches are used to determine the current patch label, using a voting scheme, thus generating the test image label map. Finally, in order to enforce spatial consistency, we iteratively incorporate a label decision from the neighboring voxels obtained during the previous iteration, by adding them to the k-NN metric. In the following subsections we provide a detailed description of the algorithm.

2.1 Creation of a labeled Patch Database

Before starting the segmentation process (in the training phase as well as for a new test image), the followed pre-processing steps were taken: brain extraction, intensity non-uniformity correction, sub-sampling and normalization.

In the training step, we extract regions-of-interest ($ROIs$) around all manually marked lesions. An ROI is defined as a bounding box of the lesion. Note that we expand the borders of the bounding box by a few voxels (three) in order to capture the surrounding tissue as well. $ROIs$ therefore represent the lesion and non-lesion areas in the brain in a roughly balanced way. From each ROI, and for each input channel, we then extract 3x3x3 patches. In this work, we treated the three modalities equally. Patches are concatenated across the input channels to form augmented patches (hereon we refer to patches as the concatenated patches). This set of patches forms the training patch database $\mathbf{D}$ (in our implementation we used 130,000 patches).

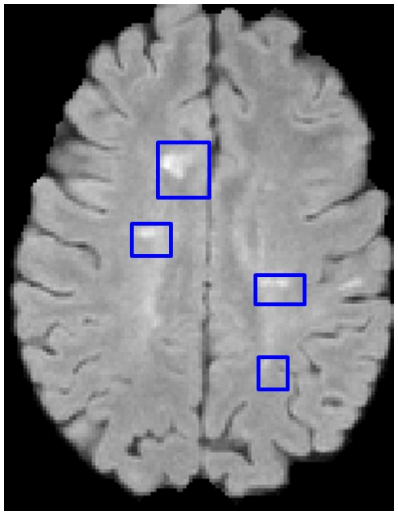

Figure 1. ROI (blue) on a FLAIR image.

2.2 Generation of an Initial Label Map

Following the procedure described above, the voxels marked by the mask are further analyzed towards classification as lesion or non-lesion using a patch-based decision method. The patch-base segmentation strategy is based on the non-local-means estimator introduced by Buades et al.[11] This method has been studied in a variety of tasks.[1,2] In this approach, similarities between multi-channel patches are used. After patches are extracted from several MR channels for a test case, similar patches are found in the labeled database described above. These labels are combined to yield a lesion segmentation map.

In the patch matching step for each voxel x in the test image, we extract a corresponding patch P_x centered at x. We then retrieve k similar patches $P_i^D(x), i = \{1 \dots k\}$ from the database $\mathbf{D}$. Each selected patch is weighted as follows:

$$w(x, P_i^D) = \exp\left(-\frac{d_I(P_x, P_i^D(x))}{\sigma^2}\right) \tag{1}$$

where $d_I(\cdot,\cdot)$ is the intensity based square Euclidean distance and σ^2 is the maximum distance measure obtained for all the voxels x and for all the patches in the database $\mathbf{D}$.

To obtain a label for a test voxel x, a vote aggregation method is defined using all labeled patches extracted via the k-NN patch matching procedure. The voxel x appears in 27 test patches which are centered on x and centered on its neighboring voxels. For each test patch we extract k nearest neighbor labeled patches from the data-set $\mathbf{D}$. Each retrieved similar patch, along with its labels and its weight, contributes a vote. Thus, for a

given voxel x we have $27 \cdot k$ votes that are averaged. The aggregation of votes results in a probabilistic lesion decision, as follows:

$$p(\text{lesion}|x) = \frac{\sum_{y \in P_x} \sum_{i=1}^{k} w(y, P_i^D(y)) \cdot l_i^D(y, x)}{\sum_{y \in P_x} \sum_{i=1}^{k} w(y, P_i^D(y))} \quad (2)$$

where y goes over the 27 neighboring voxels of x and $l_i^D(y, x)$ is the label that the training patch $P_i^D(y)$ (which is similar to the patch centered at y) assigns to x. The final lesion label, $L(x)$, for each voxel x, is determined as 1 if $p(\text{lesion}|x) > 0.5$ and 0 otherwise. This content-based segmentation process results in an initial label map.

2.3 Patch-based Label Refinement

The above procedure treats each voxel independently and ignores spatial consistency constraints. For example, a case may occur where a single voxel is labeled as non-lesion even though all its neighboring voxels are labeled as lesions. In the following, we focus on incorporating spatial consistency into the lesion map result. Spatial information was shown to improve performance.[5] For example, in[12] a refinement step based on anatomical information was included.

In our method we enforces spatial consistency using a new metric for the patch-based voting aggregation algorithm described above. The new metric uses the current label map in addition to the intensity information. A weighting parameter, α, is used to determine the relative importance of the intensity similarity and the spatial consistency constraint. This new metric is used in the patch matching step and in the vote weighing step. The metric between a test patch P_x and a training patch P_D is defined as follows:

$$d(P_x, P_D) = d_I(P_x, P_D) + \alpha \cdot d_L(P_x, P_D) \quad (3)$$

where $d_I(\cdot, \cdot)$ is the intensity based metric described above and $d_L(P_x, P_D)$ is the square Euclidean distance between the current labels of the patch P_x and the labels of training patch P_D. A larger weight parameter α, tends to favor the current labeling decision of the algorithm. We use the following definition for the weight:

$$\alpha_t = \alpha_0 * (t - 1) \quad (4)$$

with constant α_0 ensuring that the intensity metric and label metric are on the same scale and t is the iteration number. In the first iteration, $\alpha = 0$; thus we start with the patch-based method as described in Section 2.3. As α is increased from one iteration to the next, it gives more weight to labels of the nearest-neighbor training patches in making the current label decision. A large α ensures that the labels of the nearest-neighbor training patches coincide with the current label decision and therefore the algorithm converges. We have found that in practice there is no need for more than 5-6 iterations until convergence.

Our method imposes a global consistency constraint between the labels of neighboring voxels in an efficient way. For each patch we find the most similar patch based on intensity while also taking into account spatial compatibility with the neighboring voxels. The initial patch-based segmentation results are sensitive to small hyper-intense regions that are caused by noise and inhomogeneities. These regions contain patches that are similar to lesion patches. The majority of these voxels are correctly labeled as non-lesions during the refinement stage.

3. EXPERIMENTAL RESULTS

We evaluated our framework using clinical data provided by the MS lesion segmentation challenge which was introduced at the MICCAI MS lesion segmentation workshop 2008.[10] The most recent works in the field use this dataset as it provides a benchmark for algorithm comparisons. We used all cases from the University of North Carolina (UNC) dataset in the provided resolution of 1mm^3 isotropic in three different modalities (T1w, T2w and FLAIR). Our procedure was evaluated in a leave-one-out framework. We computed true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), and used the following validation measures:

- True Positive Rate: $TPR = \frac{TP}{TP+FN}$.

Table 1. TPR/PPV/DSC results (%) on clinical brain image data from the MICCAI 08 challenge.[10] Our method is compared to three state-of-the-art methods.

	Souplet[7]		**Geremia**[6]		**Weiss**[8]			**Patch-based**		
Patient	TPR	PPV	TPR	PPV	TPR	PPV	DSC	TPR	PPV	DSC
UNC01	1	1	2	1	33	29	31	0	0	0
UNC02	37	39	48	36	54	51	53	67	34	45
UNC03	12	16	24	35	64	27	38	57	25	35
UNC04	38	54	54	38	40	51	45	65	17	27
UNC05	38	8	56	19	25	10	16	36	9	14
UNC06	57	9	15	8	13	55	20	37	69	48
UNC07	27	18	76	16	44	23	30	51	48	49
UNC08	27	20	52	32	43	13	20	24	11	15
UNC09	16	43	67	36	69	6	11	29	35	32
UNC10	22	28	53	34	43	23	30	44	45	45
Average	28	24	45	26	43	29	29	**41**	**29**	**31**

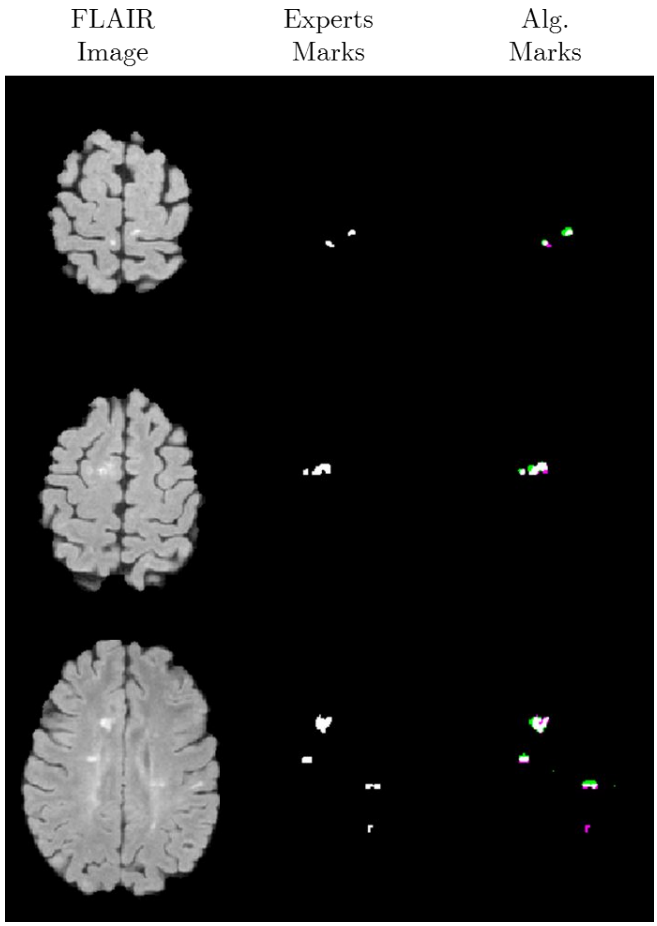

Figure 2. Comparison of alg. segmentation results and the expert's manual segmentation. Left column - FLAIR image. Middle column - expert's manual segmentation. Right column - alg. segmentation results. In the right col- TP=white, FP=green, FN=pink and TN=black.)

- Positive Predictive Value: $PPV = \frac{TP}{TP+FP}$.
- Dice Similarity Coefficient: $DSC = \frac{2 \cdot TP}{FP+FN+2 \cdot TP}$.

The measures were computed using the expert label map provided in the dataset. The scores appear in Table 1 along with a comparison to three state-of-the-art works using the same dataset. Note that DSC measures were not provided by the authors in the first two methods.

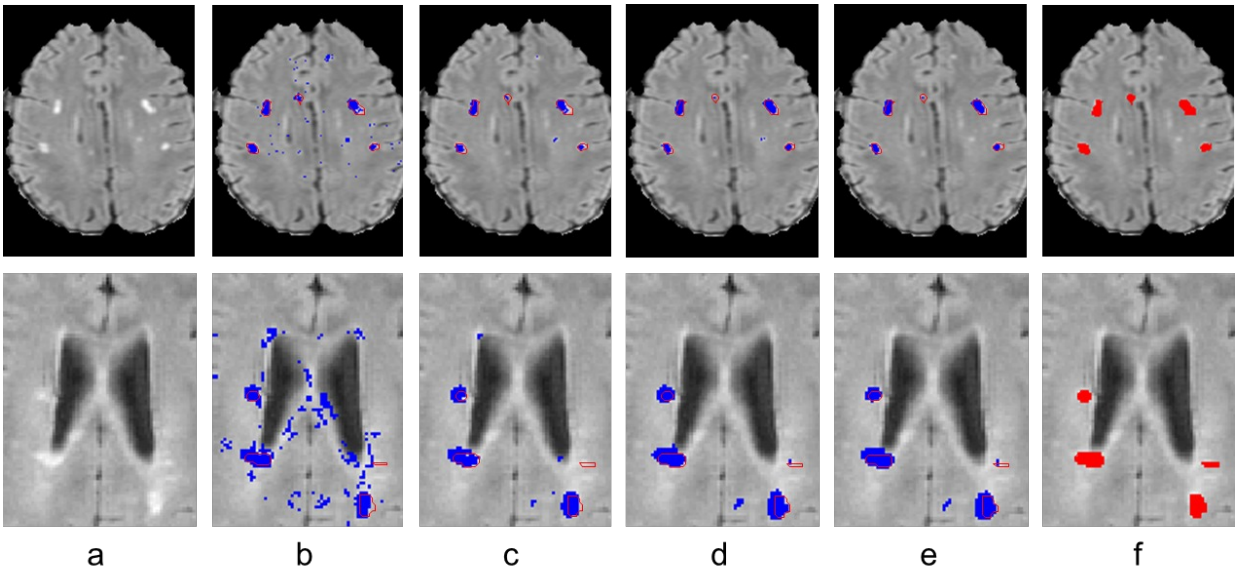

Figure 3. Two segmentation examples (top: UNC06, bottom: UNC02): (a) FLAIR image; (b) first iteration; (c-e) iterations 2,3 and 5; (f) ground truth. Proposed method in Blue; Reference segmentation in Red.

The proposed method achieved a mean TPR of 41%, a mean PPV of 29% and a mean DSC of 31%. These results are comparable with state-of-the-art algorithms. Two successful segmentation examples are shown in Fig. 3. The input image is presented in Fig. 3(a) and the final segmentation map is shown in Fig. 3(e). The similarity to the reference segmentation, shown in Fig. 3(f), is evident. We note the substantial decrease in FPs, as part of the algorithm iterative refinement procedure, in both cases. The reduction in FPs is most significant in the second iteration. We also observe an increased number of border voxels being classified as lesions in iterations 3 to 5. Thus overall we get a significant decrease in the quantity of the FPs and an increase in all the measured indices.

To conclude, we presented a novel framework for patch-based segmentation that integrates intensity information with a patch-based label refinement. The proposed framework segments the MS image without requiring registration and with no need for an atlas. This makes the framework robust to registration errors which occur, particularly if there is a high degree of anatomical variability. We applied our approach to clinical image data from the MS lesion challenge. The results are competitive with state-of-the-art methods, achieving an average DSC of 31%. This framework could be coupled with dictionary learning, as well as additional pre- and post-processing to improve performance. Future work entails improving the results using better database characterization and advanced metric learning. Evaluation of our algorithm on larger datasets from varying input sources will allow us to test the robustness of the algorithm to noise and other variance.

REFERENCES

[1] Coupé, P., Manjón, J. V., Fonov, V., Pruessner, J., Robles, M., and Collins, D. L., "Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation," *NeuroImage* **54**(2), 940–954 (2011).

[2] Rousseau, F., Habas, P. A., and Studholme, C., "A supervised patch-based approach for human brain labeling," *IEEE Tran. Medical Imaging* **30**(10), 1852–1862 (2011).

[3] Roy, S., He, Q., Carass, A., Jog, A., Cuzzocreo, J. L., Reich, D. S., Prince, J., and Pham, D., "Example based lesion segmentation," in [*SPIE Medical Imaging*], 90341Y–90341Y, International Society for Optics and Photonics (2014).

[4] Aït-Ali, L. S., Prima, S., Hellier, P., Carsin, B., Edan, G., and Barillot, C., "STREM: a robust multidimensional parametric method to segment MS lesions in MRI," in [*MICCAI*], 409–416, Springer (2005).

[5] García-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D. L., and Collins, D. L., "Review of automatic segmentation methods of MS white matter lesions on conventional magnetic resonance imaging," *Medical image analysis* **17**(1), 1–18 (2013).

[6] Geremia, E., Clatz, O., Menze, B. H., Konukoglu, E., Criminisi, A., and Ayache, N., "Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images," *NeuroImage* **57**(2), 378–390 (2011).

[7] Souplet, J.-C., Lebrun, C., Ayache, N., Malandain, G., et al., "An automatic segmentation of T2-FLAIR multiple sclerosis lesions," in [*The MIDAS Journal-MS Lesion Segmentation (MICCAI Workshop)*], (2008).

[8] Weiss, N., Rueckert, D., and Rao, A., "Multiple sclerosis lesion segmentation using dictionary learning and sparse coding," in [*MICCAI*], 735–742, Springer (2013).

[9] Elliott, C., Collins, D., Arnold, D., and Arbel, T., "Temporally consistent probabilistic detection of new multiple sclerosis lesions in brain MRI," *IEEE Trans. Medical Imaging* **32**(8), 1490–1503 (2013).

[10] Styner, M., Lee, J., Chin, B., Chin, M., Commowick, O., Tran, H., Markovic-Plese, S., Jewells, V., and Warfield, S., "3D segmentation in the clinic: A grand challenge II: MS lesion segmentation," *MIDAS Journal* , 1–5 (2008).

[11] Buades, A., Coll, B., and Morel, J.-M., "A non-local algorithm for image denoising," in [*CVPR*], **2**, 60–65, IEEE (2005).

[12] Wang, L., Shi, F., Gao, Y., Li, G., Gilmore, J. H., Lin, W., and Shen, D., "Integration of sparse multi-modality representation and anatomical constraint for isointense infant brain MR image segmentation," *NeuroImage* **89**, 152–164 (2014).