

AUTOMATED SEGMENTATION OF BRAIN LESIONS BY COMBINING INTENSITY AND SPATIAL INFORMATION

Bilwaj Gaonkar, Guray Erus, Nick Bryan, Christos Davatzikos

Section of Biomedical Image Analysis,
Department of Radiology, University of Pennsylvania, Philadelphia, PA, 19104, USA

ABSTRACT

Quantitative analysis of brain lesions in large clinical trials is becoming more and more important. We present a new automated method, that combines intensity based lesion segmentation with a false positive elimination method based on the spatial distribution of lesions. A Support Vector Regressor (SVR) is trained on expert-defined lesion masks using image histograms as features, in order to obtain an initial lesion segmentation. A lesion probability map that represents the spatial distribution of true and false positives on the intensity based segmentation is constructed using the segmented lesions and manual masks. A k-Nearest Neighbor (kNN) classifier based on the lesion probability map is applied to refine the segmentation.

Index Terms— Lesion Segmentation, Machine Learning, kNN, SVR, Spatial Learning

1. INTRODUCTION

Brain lesions, especially White Matter Lesions (WMLs), are associated with cerebrovascular disease, and also with aging. WMLs are common abnormalities of the brain, which may be the result of different brain diseases, such as multiple sclerosis and vascular dementia, or may appear in normal elderly subjects. MR imaging is widely used for diagnosing such diseases clinically. Manual lesion segmentation by trained experts, a commonly used method, is extremely time consuming, and suffers from high intra-observer and inter-observer variability. This raises the demand for automated lesion segmentation methods which can reduce both the cost of analysis and the intra-observer and inter-observer variability.

Several methods have been developed to automate the process of lesion detection [1, 2, 3, 4]. In [1] the authors use Support Vector Machines with Adaboost to learn a classifier from multi-modality MRIs. While this method can learn from expert delineated datasets, it completely ignores spatial distribution of lesions. On FLAIR images WMLs show up as hyperintensities with respect to surrounding healthy white matter (WM) tissues. However, their intensity range also overlaps with normal gray matter (GM) tissues. Furthermore, artefacts from skull stripping such as regions near the eyes, or

artefacts inside the ventricles may fall in the same intensity range as WMLs. This intensity overlap causes the failure of segmentation methods based solely on image intensity and is the major difficulty in accurately segmenting WMLs.

In [2] the intensity and the spatial information are combined together in a voxel based feature vector and a kNN classifier is used to segment lesions. We followed a similar approach in using the spatial information; however we propose a two level method, where the spatial information is used to refine an intensity based segmentation.

One major contribution of our method is that it does not require training on individual voxel intensities, unlike [1, 2], but learns to predict an adaptive threshold from the image histograms using a Support Vector Regressor (SVR), which is a robust machine learning method known to generalize well in high dimensional spaces. It effectively examines the pattern presented by the histogram, and determines a ‘dynamic’ threshold, i.e. a unique threshold for each individual.

Based on the correspondences between manually extracted lesions and the segmented hyperintensities, we construct spatial probability maps of true and false positives resulting from the intensity based segmentation. We propose a compact representation of the spatial information, by using centroids and volumes of connected components. The spatial maps are used by a kNN classifier in order to eliminate the false positives and to obtain the final segmentation.

We evaluated our method on a WML segmentation task on 38 MR scans for which expert defined truth was available. We obtained very promising results, which outperformed the ones reported in [1].

2. METHOD

We present a detailed description of the two main modules of the proposed method in the following subsections. The intensity based segmentation method is presented in section 2.1. The construction of the spatial probability maps and the kNN classifier is explained in section 2.2

2.1. Intensity based Segmentation

Brain lesions tend to be bright in FLAIR images. However, variations of image intensities of various brain tissues, and of the lesions, makes it difficult to derive a unique threshold that isolates the lesions. The premise of this paper is that a learning-based method can be trained to examine the patterns of the histogram, and determine an optimal threshold for each scan, which segments the lesions.

SVR [5] is an optimization based approach for solving machine learning regression problems, based on Support Vector Machines. It is a well established, nonlinear method that has been found to work very well in many regression problems.

A set of n training images $I = \{I_i \in \Omega\}_{i=1}^n$ defined over the image domain Ω , and the corresponding manual lesion masks $I^M = \{I_i^M\}_{i=1}^n$, in which lesions have been manually segmented by an expert, are used to learn the SVR model. We have chosen to use a Radial Basis Function kernel that empirically overperformed the linear kernel.

The image histograms $h(I_i) \in \mathbb{N}^{256}$ are used as feature vectors to train the SVR. The main reason of this choice is that we aim to learn an adaptive threshold based uniquely on the intensity information. The histogram summarizes the intensity information of the image in a low dimensional vector (relative to the dimension of the image domain).

SVR is trained to predict the threshold that optimally segments the lesions. The Jaccard ratio is used as a metric to measure the similarity of the segmented image and the manual mask. The Jaccard ratio is defined as:

$$J(I_i^t, I_i^M) = \frac{\Sigma_{\Omega}(I_i^t \cap I_i^M)}{\Sigma_{\Omega}(I_i^t \cup I_i^M)}$$

where I_i^t , is the image I_i thresholded at intensity value $t \in \{1, \dots, 256\}$. The optimal threshold t_i^* is determined as the one that maximizes $J(I_i^t, I_i^M)$

$$t_i^* = \arg \max_t J(I_i^t, I_i^M)$$

The SVR method, which is illustrated on Figure 1, learns the nonlinear map $\phi : \Omega \rightarrow \mathbb{N}$ that predicts the optimal threshold for a given image. After the SVR model is learned, each training image is segmented using the optimal threshold given by the model, in order to obtain $\{I_i^S\}_{i=1}^n$, the set of segmented training images. These images are used in the next module for learning the spatial distribution of lesions.

2.2. Lesion Classification based on Spatial Distribution

We intend to learn a spatial map which is capable of predicting the probability that a hyperintensity occurring at a particular position is a true lesion. This map is then used to construct a classifier that classifies hyperintensities as lesion or non-lesion according to their locations.

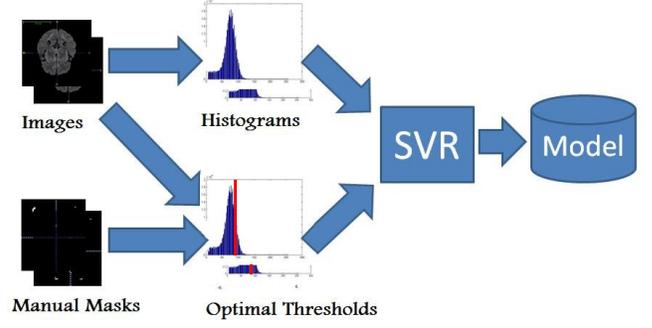


Fig. 1. Construction of the SVR model

Since we aim to learn the spatial distribution, it is necessary that all training images be in a common template space Ω_T . This is achieved approximately by linearly registering training images to Ω_T . The corresponding manual masks are also transformed to the template space.

The spatial probability maps are constructed by a voxel-wise quantification of correct detections (true positives) and misdetections (false positives) on the segmented training images $\{I_i^S\}_{i=1}^n$ according to the corresponding manual masks $\{I_i^M\}_{i=1}^n$. The binary true positive and false positive masks, I_i^{TP} and I_i^{FP} , are calculated for each training image $i = \{1 \dots n\}$.

Instead of keeping each voxel, the true positives are stored in an efficient way by calculating the centroids $\mathbf{c}^+ = \{c_x^+, c_y^+, c_z^+\}$ and the total volumes \mathbf{v}^+ of each 26-connected region in $\{I_i^{TP}\}_{i=1}^n$. In a similar way \mathbf{c}^- and \mathbf{v}^- are calculated from $\{I_i^{FP}\}_{i=1}^n$.

To each centroid a weight corresponding to its volume and its sign ($w_i^+ = v_i^+, w_i^- = -1 \cdot v_i^-$) is assigned. As the number of false positives is typically much higher than the number of true positives, the weights are normalized as follows:

$$w_i^+ = \frac{w_i^+}{\sum_{j=1}^p w_j^+}, \forall i = 1, \dots, p$$

$$w_i^- = \frac{w_i^-}{\sum_{j=1}^q w_j^-}, \forall i = 1, \dots, q$$

where p and q are respectively the number of centroids in \mathbf{c}^+ and \mathbf{c}^- , such that

$$\sum_{i=1}^p w_i^+ = 1 \quad \text{and} \quad \sum_{i=1}^q w_i^- = -1$$

A kNN classifier that aims to refine the initial lesion segmentation is constructed using $[\mathbf{c}^+; \mathbf{w}^+]$ and $[\mathbf{c}^-; \mathbf{w}^-]$ as positive and negative samples. At each voxel initially segmented as a lesion, the nearest k centroids are detected from the set

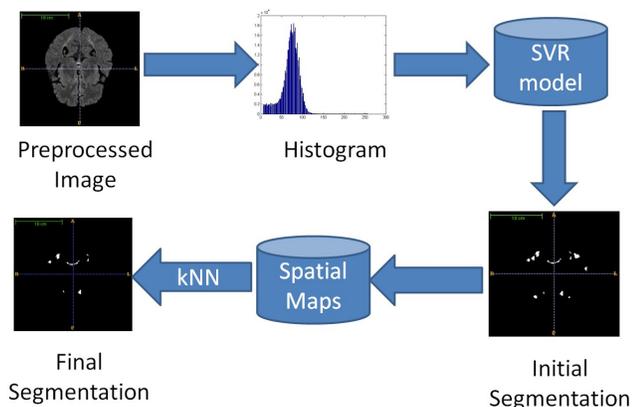


Fig. 2. Segmentation of lesions on a test image using the SVR model and the spatial maps learned from the training samples

$\{c^+, c^-\}$. The voxel is classified as a lesion if the sum of the selected centroids' weights is positive. It is classified as a false positive otherwise. Alternatively a kNN threshold different than 0 can be used for a more conservative or more aggressive false positive elimination, depending on the confidence level that we have for the intensity based segmentation in the first module.

3. EXPERIMENTAL RESULTS

We evaluated the performance of the proposed method on 38 FLAIR images manually outlined by an expert neuroradiologist. The manual lesion masks are considered as gold standard for training and evaluating prospective lesion segmentation methods.

As the number of samples was limited, cross validation by a leave-one-out strategy is applied, by training the method using 37 subjects' MRIs and applying the learned model on the left out subject, for each of the 38 subjects.

Image preprocessing on the FLAIR images involves skull stripping using the BET algorithm implemented in FMRIB's FSL software library [6], and bias correction done using N3 [7]. Furthermore, an intra-subject rigid histogram matching is applied. The images and manual masks are registered to a common template using FMRIBs linear registration tool [8].

We first train the SVR using the histograms of the training images and the manual masks. The optimal threshold for each training image is detected by applying the learned model, and the initial segmentation is performed. We then construct the spatial probability maps using the segmented images and manual masks. The lesions on a test image are detected by first segmenting it using the optimal threshold given by the SVR model, and then eliminating the false positives using kNN according to the spatial maps. Figure 2 illustrates the application of the processing pipeline on a test image for

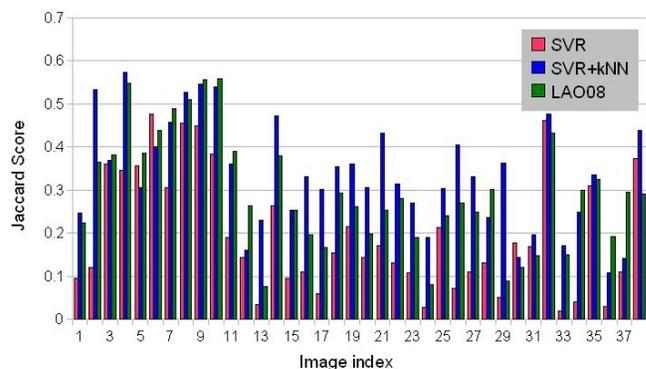


Fig. 3. Jaccard scores for the segmentation of the 38 images using our method (initial and final) and *Lao08* method

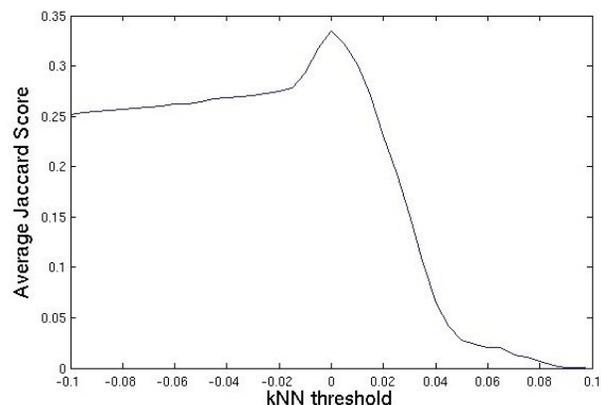


Fig. 4. Averages of the Jaccard scores of the 38 images for different values of kNN threshold

segmenting lesions.

We calculated the Jaccard score to evaluate the quality of the lesion segmentation obtained using the proposed method. Figure 3 presents the Jaccard scores of the 38 subjects both for the initial and final segmentations, compared with the scores obtained using [1] (*Lao08*). We obtained higher Jaccard scores for majority of the subjects, and a higher average score (0.3346 as opposed to 0.2928 in *Lao08*). The Jaccard score is significantly improved by the false positive elimination step for almost all subjects.

We also calculated Jaccard scores for a set of kNN thresholds in the interval $[-0.1, 0.1]$ in order to detect the threshold value that gives the highest average Jaccard score. We found that the optimal kNN threshold is 0, which shows that the sign of the kNN score can be used directly to classify false positives (Figure 4).

We compared the total lesion volume on segmented images $\{v_i^S\}_{i=1}^n$ to the lesion volume on manual masks $\{v_i^M\}_{i=1}^n$. We calculated d , the ratio of the cumulative vol-

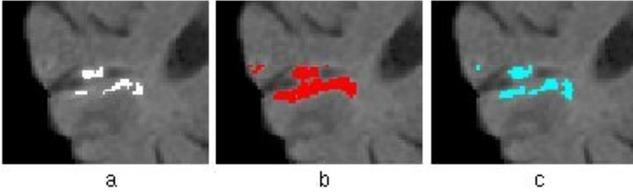


Fig. 5. Example of lesion segmentation, the masks are overlaid on the original image. a) The manual mask (white), b) the segmentation by *Lao08* (red), c) the segmentation by our method (teal).

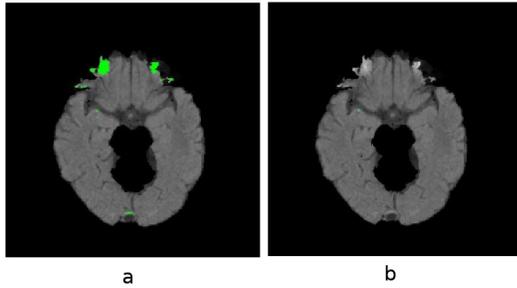


Fig. 6. Example of false positive elimination. a) The lesion mask obtained by the intensity based segmentation (green), b) the final segmentation after false positive elimination, where almost all the false positives have been eliminated (teal)

ume difference to the total lesion volume, as

$$d = \frac{\sum_{i=1}^n |v_i^S - v_i^M|}{\sum_{i=1}^n v_i^M}$$

This ratio is 0.2893 for our method and 0.8432 for *Lao08*.

Figure 5 presents an example of a lesion segmented by both methods, together with the ground truth. The boundary of the lesion segmented by our method is closer to the manually segmented lesion. We also observed that our method was successful in removing false positives that result from artefacts in the cortical regions (Figure 6).

4. CONCLUSION

We presented a new lesion segmentation method that aims to combine intensity and spatial information, in a way that is consistent with the approach used by the radiologists for manual segmentation. We proposed a two step approach where an SVR model is trained for intensity based segmentation, and kNN is used for false positive elimination according to the spatial distribution of lesions.

We obtained very promising results. A qualitative analysis of the results indicates that the SVR model is successful

in detecting an optimal threshold that adapts to the intensity profile of a given image. The false positive elimination was efficient in removing hyperintensities misdetected as lesions, based on their location. This was especially useful in removing false positives that were due to artefacts from preprocessing steps.

As a short term perspective, we believe that the method can be made more flexible by replacing the hard-thresholding in the first stage by a fuzzy segmentation, and adapting the second stage to the resulting fuzzy labels.

Because we do not train our models on voxelwise attribute vectors, training is very fast. Consequently, the method can be easily adapted to online learning and be deployed in a clinical setting.

5. REFERENCES

- [1] Z. Lao, D. Shen, D. Liu, A. F. Jawad, E. R. Melhem, L. J. Launer, R. N. Bryan, and C. Davatzikos, "Computer-assisted segmentation of white matter lesions in 3d mr images using support vector machine," *Academic Radiology*, vol. 15 No 3, pp. 300–313, 2008.
- [2] P. Anbeek, K. L. Vincken, M.J.P. van Osch, R.H.C. Bisschops, and J. van der Grond, "Automatic segmentation of different-sized white matter lesions by voxel probability estimation," *Medical Image Analysis*, vol. 8, no. 3, pp. 205 – 215, 2004.
- [3] Van Leemput K., F. Maes, D. Vandermeulen, Colchester A., and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Transactions on Medical Imaging*, vol. 20(8), pp. 677 – 688, 2001.
- [4] M. Prastawa and G. Gerig, "Brain lesion segmentation through physical model estimation," in *ISVC (1)*, 2008, pp. 562–571.
- [5] Alex J. Smola and Bernhard Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [6] S.M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping*, vol. 17, pp. 143–155, 2002.
- [7] A. Evans J. Sled, A. Zijdenbos, "A nonparametric method for automatic correction of intensity nonuniformity in mri data.," *IEEE Trans Med Imaging*, vol. 17, pp. 87–97, 1998.
- [8] S.M. Smith, M.Jenkinson, and M.W. Woolrich, "Advances in functional and structural mr image analysis and implementation as fsl.," *Neuro-Image*, vol. 23, pp. 208–219, 2004.