# Minimization of False MS Lesion Classifications on MR Images: Quantitative Validation

B. R. Sajja[1], S. Datta[1], R. He[1], P. A. Narayana[1]

[1]Diagnostic and Interventional Imaging, University of Texas Medical School at Houston, Houston, Texas, United States

**Introduction:** Accurate segmentation is a prerequisite for robust estimation of multiple sclerosis (MS) lesion volumes. Unfortunately, results obtained from most of the segmentation methods are greatly affected by various factors, such as noise, intensity inhomogeneities, and partial volume effects. The presence of large number of false classifications requires manual intervention for validating the lesion classifications. This introduces significant operator bias. More importantly such intervention is impractical when dealing with large number of images that are typically encountered in multicenter clinical trials. In this study a unified approach that combines both nonparametric (Parzen window classifier [1]) and parametric (HMRF-EM [2]) methods along with morphological operations is presented for minimizing the false lesion classifications. This approach is validated in 23 MS patients.

**Methods and Materials:** MR images on 23 clinically definite MS patients were acquired on a GE 1.5 T MR scanner. The protocol includes the acquisition of dual Fast Spin Echo (FSE) and Fluid Attenuation Inversion Recovery (FLAIR) images with slice thickness of 3 mm, FOV of 240 mm x 240 mm, and image matrix of 256x256. The preprocessing steps, prior to the segmentation procedure, include registration of FLAIR images with FSE images, bias field correction, anisotropic diffusion filtering, brain stripping and intensity standardization. Initially, lesions, CSF and parenchyma were segmented on late-echo FSE and FLAIR images based on the feature map obtained using Parzen window algorithm [1]. The false positive lesion classifications were eliminated using the ratio map of PD and T2 weighted images. A single threshold value for all data sets was used to obtain a binary ratio mask of brain, excluding lesions and CSF. Negation of this mask was multiplied by the lesions obtained with Parzen classifier. Since the lesion boundary on PD and T2 weighted images is generally more diffuse compared to FLAIR image, a region growing algorithm was used to obtain complete lesion size shown on the initial feature map classification. Now the remaining brain parenchyma excluding lesions and CSF is classified into gray matter (GM) and white matter (WM) using PD and T2 weighted images utilizing HMRF-EM algorithm, a parametric technique [2]. This method gives better GM and WM classification by exploiting the contextual information through hidden Markov random field. During the minimization of false positives, some of the true lesions, particularly subtle lesions, were eliminated from the lesion classification. After GM and WM segmentation these regions were mostly classified as GM. More than 95% of the MR-visible MS lesions reside in the white matter. So, all the GM islands surrounded by WM or lesion were identified using mathematical morphological operations and revisited for lesion classification. These GM regions were verified for lesions on the earlier Parzen window classification. After minimizing the false classifications, lesions were delineated using the fuzzy connectivity on the FLAIR images. As judged by an expert neuroradiologist, the final segmentation results have shown more accurate lesion sizes with significantly reduced false lesion classifications.

Evaluation was performed by quantitatively comparing the segmentation results with those manually outlined by an expert neuroradiologist, considered to be the gold standard or reference. Depending on the manual segmented lesion volume (LV) in the entire brain, all 23 patients were classified into two categories with Category I: LV < 10 cc and Category II: LV >= 10 cc. Out of 23 patients, 8 were in category I and 15 were in category II. The binary segmentations were evaluated by four different similarity measures: similarity index (SI), percentage of correct estimation (PCE), percentage of over estimation (POE), and percentage of under estimation (PUE). The SI is a measure for the correctly classified lesion volume relative to the total volume of lesion in both the reference and the segmented image. The PCE measures the percentage of correctly classified lesion volume relative to the lesion volume in the reference. The POE measures the percentage of false positive classification relative to the lesion volume in the reference while PUE measures the percentage of missed lesion classification. The similarity measures are defined as:

$$\mathrm{SI} = \frac{2\times(\mathrm{Ref}\cap\mathrm{Seg})}{\mathrm{Ref}+\mathrm{Seg}},\quad \mathrm{PCE} = \frac{\mathrm{Ref}\cap\mathrm{Seg}}{\mathrm{Ref}}\times 100,\quad \mathrm{POE} = \frac{\overline{\mathrm{Ref}}\cap\mathrm{Seg}}{\mathrm{Ref}}\times 100,\quad \mathrm{PUE} = \frac{\mathrm{Ref}\cap\overline{\mathrm{Seg}}}{\mathrm{Ref}}\times 100.$$

Here, Ref and Seg denote the lesion volumes of reference and automatic segmentation respectively. $\mathrm{Ref}\cap\mathrm{Seg}$ represents the volume of the correctly classified voxels. The volume $\overline{\mathrm{Ref}}\cap\mathrm{Seg}$ corresponds to the false positives. Similarly, $\mathrm{Ref}\cap\overline{\mathrm{Seg}}$ represents the false negatives [3].

**Results and Discussion:** Figure 1 shows the similarity indices for Parzen classification of lesions and following the application of the false positive minimization (FPM) and false negative minimization (FNM) steps for the category I patients. The increase in SI following FNM indicates the good agreement of segmented lesions with manual segmentation. Figures 2 and 3 show the quantitative measure of POE and PUE for Parzen, FPM, and FNM cases. The plots look similar for the category II subjects except for larger sd. The quantitative measures for all subjects are shown in Table 1 (mean ± sd). The large errors for category II compared to category I are due to the fact that small errors in small lesions have a relatively large effect on the similarity measures [3]. The same feature map and the thresholds were used for segmenting all the images. Thus the proposed method is fully automatic, except for limited human intervention in stripping the images.

**Conclusion:** A method for automatically identifying and eliminating both false positive and false negative lesion classifications on MRI of MS brains has been presented. This method combined Parzen window classifier and HMRF-EM method along with intermediate image processing and mathematical morphological steps. The validation results showed that this method produced excellent lesion classification.

**Acknowledgement:** This work was supported by NIH Grant RO1 EB002095.

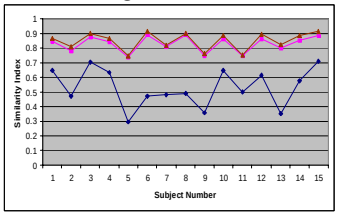


Figure 1

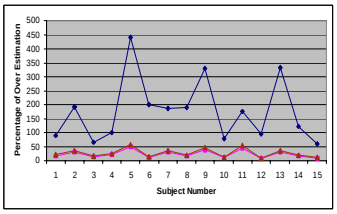


Figure 2

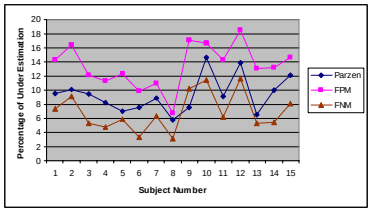


Figure 3

Plots of SI (Figure 1), POE (Figure 2), and PUE (Figure 3) for lesion volumes at Parzen, FPM, and FNM steps for category I subjects.

Table 1: Similarity measures for category I, category II and all subjects.

| Category | POE | PUE | PCE | SI |
|---|---|---|---|---|
| I (8) | 59.36±34.34 | 19.74±20.29 | 80.26±20.29 | 0.67±0.14 |
| II (15) | 27.19±15.98 | 6.94±2.70 | 93.06±2.70 | 0.85±0.06 |
| I+II (23) | 38.38±27.98 | 11.38±13.22 | 88.62±13.22 | 0.79±0.13 |

**References:** [1] Duda RO, Hart PE, Stark DG. New York, John Wiley & Sons; 2001. [2] Zhang Y et al. IEEE Trans Med Imag. 2001;20:45-57. [3] Anbeek P et al. NeuroImage 2004;21:1037-104.