# Automated lesion detection from multimodal brain MRI using Markov random fields and random forest

Jhimli Mitra[1], Soumya Ghose[1], Pierrick Bourgeat[1], Olivier Salvado[1], Stephen Rose[1], Bruce Campbell[2], Alan Connelly[3], Susan Palmer[4], Leeanne Carey[4], and Jurgen Fripp[1]

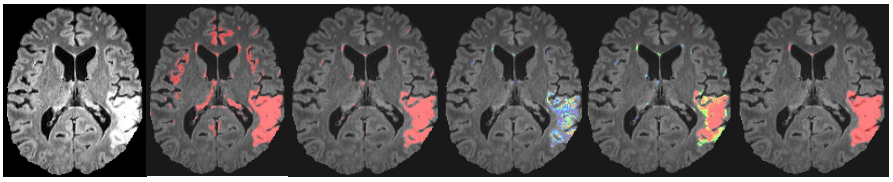
*[1]The Australian e-Health Research Centre, CSIRO Computational Informatics, CSIRO Preventative-Health Flagship, Herston, QLD, Australia, [2]Department of Medicine and Neurology, Royal Melbourne Hospital, Parkville, VIC, Australia, [3]Department of Radiology, Royal Melbourne Hospital, Parkville, VIC, Australia, [4]The Florey Institute of Neuroscience and Mental Health, Parkville, VIC, Australia*

**Introduction:** Pathological changes occur within the ischemic lesion over time post-stroke, and therefore, MRI tissue contrasts are often time-dependent and heterogeneous in nature; i.e., they vary between acute (less than 7 days) and chronic (>=3 months) stages[1]. Understanding structure-function relationships in the brain after stroke is reliant not only on the accurate anatomical delineation of the focal ischemic lesion, but also on previous infarcts, remote changes and the presence of white matter hyperintensities (WMH) [2, 3]. In the acute stage, hyperintense signal observed on diffusion-weighted images (DWI) provides important information about the anatomical location and extent of the infarcted territory, while at a more chronic stage T2W or fluid-attenuated inversion recovery (FLAIR) MRI are used to delineate the stroke lesion volume. Despite the probable differences in the underlying tissue pathologies of ischemic strokes and WM lesions, the tissue textures appear similar in MR images, i.e., they are hyperintense in T2W and FLAIR and hypointense in T1W images. The complexity of delineation further increases with the location of the infarct, for e.g., when the ischemic lesion occurs within deep WM regions that are anatomically coincident with the presence of WMH. Distinguishing between the two types of lesions is extremely difficult using FLAIR, and even more so in the absence of acute DWI images[4].

In this paper, we present an automated method to segment the chronic ischemic stroke, white-matter and silent brain infarcts in a two-stage classification scheme involving maximum aposteriori-Markov random field (MAP-MRF) and random forest, from FLAIR, T1W, T2W and apparent diffusion coefficient (ADC) maps and in the absence of acute stage DWI images.

**Method:** Our study involves a cohort of 36 stroke patients from the START[5] study. The MR images were acquired using a 3T MR (Magnetom Trio; Siemens, Erlangen, Germany) scanner. The T1W images were acquired with MPRAGE sequence with TE=2.55ms, TR=1900ms, flip angle=9˚, isotropic voxel resolution=1mm. The T2W FLAIR images were acquired with TE=388ms, TR=6000ms, T1=2100ms, flip angle=120˚ with voxel resolution=1mm×0.5mm×0.5mm. The ADC maps were generated from DWI images acquired using a spin echo diffusion-weighted echoplanar imaging (EPI) sequence with B=1000 employing 25 diffusion encoding directions with an isotropic voxel resolution of 2.5mm. Lesion volumes were manually segmented by an expert neurologist on multiplanar re-projections of the FLAIR scans with voxel resolution of 3mm×0.5mm×0.5mm. The T1W images were rigidly registered to the (Montreal Neurological Institute) MNI atlas (1mm isotropic), and all other modalities were rigidly registered to the T1W in MNI space. The T1W images were used to obtain segmented grey matter (GM) and white matter (WM) masks, while the T2W images were used to extract the cerebrospinal fluid (CSF) masks[6]. Skull-stripping was done by masking the brain with the combination of the GM, WM and CSF masks. An N4 magnetic bias-correction was done on the FLAIR.

The first stage of our method involves extracting probable candidate lesions from the hyperintense regions of the FLAIR MRI (Fig.1(a)). A MAP-MRF, which is an MRF segmentation[7] on the estimated labels after a Bayesian classification; was applied on the FLAIR to extract 5 classes as background, GM, WM, CSF and a possible lesion class. The possible FLAIR lesion class (Fig.1(b)) was further classified into 8-classes to best separate the lesions from the contaminated GM tissues. The 3-classes with high means were then merged to provide the probable FLAIR lesion class (Fig.1(c)). While the MAP-MRF provided a hard classification of the probable lesion class, a probabilistic classification was required to be attributed as a feature in the subsequent random forest training and classification. Therefore an 8-class Gaussian expectation –maximization[8] was further applied on the merged lesion class to best separate the lesion and non-lesion classes. The probabilities of the class with the highest mean were considered as lesion probabilities from the FLAIR images (Fig.1(d)).



(a) (b) (c) (d) (e) (f)

Fig.1. Effect of each stage of the proposed segmentation method (a) FLAIR image; (b) & (c) Possible lesions classes overlaid after 5- and 8-class MAP-MRF; (d) & (e) probabilities after 8-class EM and probabilities from RF (color range: blue=0, red=1); (f) Final binary classification of RF probabilities.

The random forest[9] training involved, building a forest of 100 trees with the context-aware features[10]. The feature set included variables of the probable lesion class from the multichannel MRI (T1W, T2W, FLAIR, ADC, lesion probabilities, GM and WM probabilities respectively). During random forest classification, probabilistic labels of the lesion class were obtained (Fig.1(e)) and a flat threshold of 0.2 was applied to obtain binary lesion volumes (Fig.1(f)). A *k*-fold validation method was conducted on the 36 patients with *k*=2. The MAP-MRF classification required 15 mins for each subject; the random forest training required approx 4 hrs, while each testing was done in less than 1.5 mins on a 6-core CPU of 3.2GHz with 23.5GB of memory. This did not however include feature-extraction time.

**Results:** Standard evaluation metrics were used to evaluate our automated segmentation method. Dice similarity coefficient (DSC) was used to measure the overlap ration between the automatically and manually segmented lesion volumes. Sensitivity or the recall rate was used to measure the proportion of the actual positives which were correctly identified as such, while the positive predictive value (PPV) or the precision rate was used to measure the proportion of positive test results that are true positives. Table 1 shows the mean±std. dev. values for each of these metrics measured on the stroke cohort. The Pearson's correlation between the manually and automatically segmented volumes is shown in the last column. Fig. 2 shows some of the qualitative segmentation results using our method.

Table 1. Quantitative evaluation of the segmentation method.

| Patients | DSC | Sensitivity | PPV | Volume correlation |
|---|---|---|---|---|
| 36 | 0.60±0.12 | 0.52±0.14 | 0.76±0.17 | 0.76 (p<0.00001) |

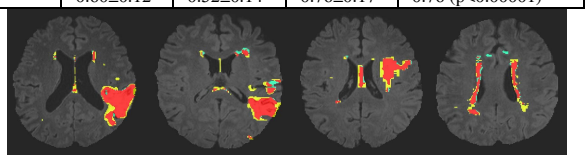


Fig.2. Qualitative segmentation results using the automated method. Color codes for overlap with the manual segmentation: red=true positives, yellow=false negatives, green=false positives.

**Discussions:** From our experiments we found that our method was successful in segmenting the chronic ischemic stroke, WMH and other infarcts that could influence the structure-function relationships of the brain post-stroke. The highest DSC obtained was 0.81 with the mean and median values were at 0.60. The DSC, sensitivity and PPV values indicate that our method performed comparably with one of the methods in literature that segmented any type of brain lesion but was evaluated on a different and limited cohort of 10 patients[11]. Since our method initially relied on FLAIR hyperintensities at the first stage to detect probable lesion regions; tissue loss due to post-stroke liquefactive necrosis (hypointense on FLAIR) could not be estimated using our method.

**References:** **1.** Carey L, et al. Beyond the lesion - Neuroimaging foundations for poststroke recovery. *Future Neurology*. 2013;8:507-524. **2.** Vermeer S.E, et al. Silent brain infarcts and white matter lesions increase stroke risk in the general population: The Rotterdam scan study. *Stroke*. 2003;34:1126-29. **3.** Norrving B. Leucoaraiosis and silent subcortical infarcts. *Reviews in Neurology (Paris)*. 2008;164:801-804. **4.** Campbell B.C, et al. Assessing response to stroke thrombolysis: validation of 24-hour multimodal magnetic resonance imaging. *Arch Neurology*. 2012;69:46-50. **5.** The START (**ST**roke im**A**ging p**R**evention and **T**reatment) study, https://www.START.csiro.au. **6.** van Leemput K, et al. Automated model-based tissue classification of MR images of the brain. *IEEE Trans in Medical Imaging*. 1999;18:897–908. **7.** Besag J. On the statistical analysis of dirty pictures. *J of Royal Stats Soc, Ser B*. 1986;48:259-308. **8.** Dempster A.P, et al. Maximum likelihood from incomplete data via the EM algorithm. *J of Royal Stats Soc, Ser B*. 1977;39:1-38. **9.** Criminisi A, et al. Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning. *TechReport MSR-TR-2011-114*, Microsoft Research, 2011. **10.** Zikic D, et al. Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR. In proc. of MICCAI, *LNCS* 75. 2012;12:369-376. **11.** Seghier M.L, et al. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *NeuroImage*. 2008;41:1253-1266.