Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors
Ana Sanjuán, Cathy J. Price, Laura Mancini, Goulven Josse, Alice Grogan, Adam K. Yamamoto, Sharon Geva, Alex P. Leff, Tarek A. Yousry, Mohamed L. Seghier

To cite this version:
Ana Sanjuán, Cathy J. Price, Laura Mancini, Goulven Josse, Alice Grogan, et al.. Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors. Frontiers in Neuroscience, Frontiers, 2013, 7, pp.241. 10.3389/fnins.2013.00241. hal-01578520

HAL Id: hal-01578520
https://hal.sorbonne-universite.fr/hal-01578520
Submitted on 29 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors

Ana Sanjuán1,2*, Cathy J. Price1, Laura Mancini3, Goulven Josse4, Alice Grogan1, Adam K. Yamamoto5, Sharon Geva5, Alex P. Leff1,6, Tarek A. Yousry3 and Mohamed L. Seghier1

1 Wellcome Trust Centre for Neuroimaging, University College of London, London, UK
2 Departamento de Psicología Básica, Clínica y Psicobiología, Universitat Jaume I, Castellón, Spain
3 Lysholm Department of Neuroradiology, National Hospital for Neurology and Neurosurgery, London, UK
4 Hôpital de la Pitié-Salpêtrière, Institut du Cerveau et de la Moëlle épinière, Paris, France
5 Developmental Cognitive Neuroscience Unit, Institute of Child Health, University College of London, London, UK
6 Institute of Cognitive Neuroscience, University College of London, London, UK

INTRODUCTION
Meaningful lesion-behavior mapping in patients with focal brain damage (Bates et al., 2003; Rorden et al., 2009) rests upon precise identification of each patient’s lesion. The most commonly used method for lesion identification is manual tracing of abnormal brain tissue by trained experts; this is conventionally taken to be the gold-standard method. However, the manual tracing of lesions is laborious, time-consuming and largely operator-dependent (Fiez et al., 2000). To overcome these limitations, several automated procedures for lesion identification have been proposed (see review in Gondal and Khan, 2013; Gordillo et al., 2013), but they generally involve multi-spectral brain images (e.g., Soltanian-Zadeh et al., 1998; Capelle et al., 2004; Prastawa et al., 2004; Wu et al., 2006b; Kabir et al., 2007; Ruan et al., 2007; Corso et al., 2008; Nie et al., 2009; Menze et al., 2010; Hsieh et al., 2011; Popuri et al., 2012). The acquisition of 3D multi-spectral anatomical images is not always possible due to financial or time constraints. Indeed, it is not unusual that only one anatomical image (usually a T1-weighted image) is acquired from patients participating in research fMRI studies, particularly when total scanning time over fMRI runs exceeds 60 min. In these cases, tumor identification needs to be based on the anatomical images available. The aim of the present study is to demonstrate the usefulness of an automated method for identifying brain tumors from single T1-weighted images.

Brain tumors can have different shapes or locations, making their identification very challenging. In functional MRI, it is not unusual that patients have only one anatomical image due to time and financial constraints. Here, we provide a modified automatic lesion identification (ALI) procedure which enables brain tumor identification from single MR images. Our method rests on (A) a modified segmentation-normalization procedure with an explicit “extra prior” for the tumor and (B) an outlier detection procedure for abnormal voxel (i.e., tumor) classification. To minimize tissue misclassification, the segmentation-normalization procedure requires prior information of the tumor location and extent. We therefore propose that ALI is run iteratively so that the output of Step B is used as a patient-specific prior in Step A. We test this procedure on real T1-weighted images from 18 patients, and the results were validated in comparison to two independent observers’ manual tracings. The automated procedure identified the tumors successfully with an excellent agreement with the manual segmentation (area under the ROC curve = 0.97 ± 0.03). The proposed procedure increases the flexibility and robustness of the ALI tool and will be particularly useful for lesion-behavior mapping studies, or when lesion identification and/or spatial normalization are problematic.

Keywords: automatic lesion identification, segmentation, spatial normalization, fuzzy clustering, MRI
FIGURE 1 | Illustration of ALI procedure. The method identifies lesion in two core procedures. (A) A modified segmentation-normalization that identified the likelihood that each voxel belonged to one of four different tissue types and (B) a fuzzy clustering algorithm that identified voxels where the tissue probability maps differed from those in healthy controls. The red arrow shows the recursive nature of the procedure for the tumor identification. The first iteration of steps (A) and (B) (1st ALI) generated a first approximation of the tumor site. The second iteration of steps (A) and (B) (2nd ALI) included the approximate lesion definition (see red arrow) as a patient-specific prior to improve the tissue segmentation. (w, normalized; sw, smoothed and normalized).

not developed specifically for brain tumors (e.g., multiple sclerosis infarct lesions or cortical malformations, see Shen et al., 2008, 2010; Thesen et al., 2011; Lladó et al., 2012), operate on 2D slices only (Xu and Mandal, 2012), or require contrast enhanced images that are not commonly acquired in research studies (Assefa et al., 2010; Harati et al., 2011; Asman et al., 2013).

Brain tumor identification can be challenging for tissue segmentation procedures. Apart from their diversity in shape, size, and location, the presence of a tumor in brain images can be associated with a highly heterogeneous and diffuse signal distribution that can be confounded by a similar range of signal intensity in the neighboring intact tissue (Velož et al., 2011). Indeed, our initial analysis using the original implementation of ALI (Seghier et al., 2008) was able to identify the correct location of the tumors but was not always able to delineate their full extent. This was driven by misclassification of atypical tissue (i.e., tumor) as gray or white matter during the segmentation step, despite the inclusion of an extra tissue class prior in the segmentation procedure (see Figure 2, top panel). To enhance its accuracy, we propose here a practical solution that uses ALI in a recursive mode: a first run of ALI provides an approximation of the tumor size and location, and a second run of ALI includes that approximation as a patient-specific “prior” to model the extra tissue class (See Figure 2, bottom panel). This practical solution that incorporates a refined patient-specific tissue class prior in the final segmentation of the patient’s brain image, significantly improved the accuracy of tumor identification. Below, we illustrate how the automated procedure deals with real lesions, using datasets from a heterogeneous sample of 18 patients with brain tumors. The automatically identified tumors were compared to manual segmentations delineated by two independent observers.

MATERIALS AND METHODS
The study was approved by the London Queen Square Research Ethics Committee. All subjects gave written informed consent prior to scanning.

DATA SETS
T1-weighted images of 18 patients (8 females, mean age: 37.3 years, age range: 16–66 years) with brain tumors, acquired as part of an fMRI language presurgical evaluation, and 64 neurologically normal subjects (age range: 21–75 years) were included in this study. The heterogenous sample of patients consisted of 9 patients with low grade gliomas (World Health Organization (WHO) grade I: 1 dysembryoblastic neuroepithelial tumor; WHO grade II: 2 oligoastrocytomas, 3 oligodendrogliomas, 1 astrocytoma and 2 diffuse astrocytomas), 8 patients with high grade gliomas (WHO grade III: 3 anaplastic astrocytomas, 1 glioneuronal tumor with neuropil-like islands, 1 diffuse astrocytoma with anaplastic progression; WHO grade IV: 3 glioblastomas) and one meningioma. Tumors varied in size and location, in both hemispheres, as listed in Table 1. Overall, tumors were predominately located in language and motor regions, with lesion size varying between 1.5 and 23 cm³ (Table 1). It is worth noting that multispectral data are typically used for tumor segmentation, in particular in the clinical setting (Gordillo et al., 2013), namely to
provide complementary information about the different tumor parts; however, we do not have multispectral data in all our 18 patients.

MRI DATA ACQUISITION
Structural data consisted of high-resolution T1-weighted MRI scans acquired on a Siemens 1.5 T Sonata scanner (Siemens Medical Systems, Erlangen, Germany) using a 3D modified driven equilibrium Fourier transform sequence (Deichmann et al., 2004), consisting of 176 sagittal partitions with an image matrix of $256 \times 224$ and a final isotropic resolution of $1 \text{ mm}^3$ (TR/TE/inversion time = 12.24/3.56/530 ms).

LESION IDENTIFICATION PROCEDURE
The present paper uses ALI in the specific context of brain tumor identification. Below we explain how ALI works and how it can be used, in a recursive mode, to refine the definition of each patient’s tumor in an informed way.

The automated lesion identification (ALI) method
ALI was initially implemented as a toolbox in the SPM software package (Wellcome Trust Centre for Neuroimaging, London, UK) to specifically deal with lesion identification on mono-spectral MRI scans. It comprises (A) segmentation-normalization and (B) outlier detection in the following steps (see Figure 1, cf. Seghier et al., 2008):

(A) A modified implementation of the unified segmentation-normalization procedure of SPM as described in Ashburner and Friston (2005). It combines segmentation, bias correction and spatial normalization through the inversion of a single unified model. Its algorithm is based on Bayesian statistics that allow previous knowledge (i.e., priors) to be incorporated. In the context of patients with brain lesions, the standard unified segmentation-normalization tends to misclassify the abnormal tissue as “normal” gray matter (GM), white matter (WM), or cerebral spinal fluid (CSF) because there is no explicit model of what constitutes
Table 1 | Relevant lesion information (type, size, and tumor location) and similarity measures (Dice and AUC).

| P   | Type lesion | Average volume [mm³] | Location | Dice between observers | Dice O1 | Dice O2 | AUC 1stALI O1 | AUC 1stALI O2 | AUC 2ndALI O1 | AUC 2ndALI O2 | Sensitivity 2ndALI O1 | Sensitivity 2ndALI O2 | Specificity 2ndALI O1 | Specificity 2ndALI O2 |
|-----|-------------|----------------------|----------|------------------------|---------|---------|---------------|---------------|---------------|---------------|---------------------|---------------------|---------------------|---------------------|
| 01  | LGG         | 15763.5              | Left insula | 0.875                  | 0.836   | 0.820   | 0.907         | 0.890         | 0.984         | 0.979         | 0.755               | 0.774               | 0.997               | 0.993               |
| 02  | HGG         | 23200.5              | Left insula | 0.918                  | 0.781   | 0.784   | 0.945         | 0.957         | 0.967         | 0.974         | 0.767               | 0.814               | 0.981               | 0.977               |
| 03  | LGG         | 4194.0               | Inter-hemispheric (SMA) | 0.581                  | 0.699   | 0.719   | 0.843         | 0.966         | 0.952         | 0.988         | 0.548               | 0.799               | 1.000               | 0.996               |
| 04  | LGG         | 4380.5               | Right precentral | 0.942                  | 0.778   | 0.782   | 0.982         | 0.986         | 0.991         | 0.991         | 0.674               | 0.675               | 0.999               | 0.999               |
| 05  | LGG         | 13541.5              | Left parietal | 0.914                  | 0.794   | 0.786   | 0.971         | 0.959         | 0.981         | 0.970         | 0.704               | 0.679               | 0.996               | 0.997               |
| 06  | LGG         | 14289.0              | Inter-hemispheric (SMA) | 0.885                  | 0.754   | 0.702   | 0.864         | 0.821         | 0.970         | 0.947         | 0.632               | 0.560               | 0.998               | 0.998               |
| 07  | LGG         | 3734.5               | Inter-hemispheric (SMA) | 0.725                  | 0.669   | 0.747   | 0.882         | 0.913         | 0.958         | 0.972         | 0.526               | 0.679               | 0.999               | 0.998               |
| 08  | LGG         | 2490.0               | Left middle frontal | 0.902                  | 0.697   | 0.699   | 0.905         | 0.912         | 0.979         | 0.979         | 0.557               | 0.563               | 1.000               | 1.000               |
| 09  | LGG         | 1586.5               | Left temporo-parietal | 0.851                  | 0.708   | 0.708   | 0.973         | 0.974         | 0.985         | 0.989         | 0.618               | 0.653               | 0.999               | 0.999               |
| 10  | LGG         | 4842.0               | Left precentral | 0.883                  | 0.788   | 0.769   | 0.908         | 0.909         | 0.983         | 0.982         | 0.659               | 0.660               | 0.999               | 0.999               |
| 11  | LGG         | 10210.0              | Left temporo-parietal | 0.912                  | 0.782   | 0.802   | 0.964         | 0.971         | 0.978         | 0.982         | 0.675               | 0.714               | 0.998               | 0.998               |
| 12  | LGG         | 11094.5              | Left insula | 0.911                  | 0.764   | 0.785   | 0.906         | 0.922         | 0.964         | 0.972         | 0.640               | 0.686               | 0.998               | 0.997               |
| 13  | LGG         | 1760.5               | Right precentral | 0.904                  | 0.592   | 0.585   | 0.973         | 0.962         | 0.987         | 0.971         | 0.508               | 0.501               | 0.999               | 0.999               |
| 14  | LGG         | 6756.0               | Right temporo-parietal | 0.771                  | 0.632   | 0.649   | 0.880         | 0.888         | 0.941         | 0.941         | 0.527               | 0.575               | 0.996               | 0.995               |
| 15  | LGG         | 2372.0               | Inter-hemispheric (SMA) | 0.717                  | 0.614   | 0.571   | 0.950         | 0.960         | 0.974         | 0.982         | 0.515               | 0.568               | 0.998               | 0.997               |
| 16  | LGG         | 7973.5               | Left insula | 0.911                  | 0.815   | 0.827   | 0.864         | 0.871         | 0.984         | 0.987         | 0.716               | 0.739               | 0.999               | 0.999               |
| 17  | LGG         | 7836.5               | Inter-hemispheric (SMA) | 0.878                  | 0.774   | 0.810   | 0.879         | 0.906         | 0.936         | 0.955         | 0.662               | 0.736               | 0.998               | 0.998               |
| 18  | LGG         | 9427.5               | Left temporo-parietal | 0.606                  | 0.472   | 0.368   | 0.816         | 0.892         | 0.878         | 0.914         | 0.333               | 0.336               | 0.996               | 0.998               |
| Average | 8080.7 | 0.838               | 0.718   | 0.717   | 0.911   | 0.925   | 0.966         | 0.971         | 0.612         | 0.651         | 0.997               | 0.997               |                  |                  |
| SD  | 58279      | 0.110               | 0.093   | 0.115   | 0.051   | 0.044   | 0.027         | 0.020         | 0.108         | 0.117         | 0.004               | 0.005               |                  |                  |

Dice’s similarity index is shown for the comparison between (1) the manually-segmented tumors of the observers and (2) each binary lesion map obtained with the recursive ALI (i.e., fuzzy sets at a U threshold of 0.2) and the manually-segmented tumors of each observer. The area under the curve (AUC) was obtained for the comparison of the manually-segmented tumors of each observer and the binary mask obtained for different U thresholds from both the standard (1st) and recursive (2nd) ALI approaches. Finally, the specificity and sensitivity of the recursive ALI (at a U threshold of 0.2) are also reported. (P, Patient; LGG, Low Grade Glioma, HGG, High Grade Glioma; M, Meningioma; O1, Observer1; O2, Observer2 and SMA, Supplementary Motor Area).
images. This fuzzy set can be thresholded at a given $U$-value to generate a 3D-binary image of the lesion. Given the difficulty in defining an objective universal threshold that works on any dataset, we previously recommend that $U$ thresholds are typically set at around 0.3 to provide accurate lesion delineation when ALI is used on images from chronic stroke patients (Seghier et al., 2008).

Below is a succinct description of the outlier detection algorithm used in ALI. The core of its algorithm is a non-iterative Fuzzy Clustering with fixed Prototypes (FCP) method previously introduced by Seghier and colleagues (2007). In contrast to standard data-driven fuzzy clustering techniques, FCP is a hypothesis-driven technique where the number and center of the clusters (i.e., prototypes) are predefined (i.e., number of clusters = number of subjects). By assuming that a lesioned brain is an outlier in relation to normal (control) brains, FCP can look for voxels that are very different in the lesioned brain as compared to controls.

The algorithm of FCP exploits the well-known fact that regression analysis is sensitive to outlier values. First, the centroids of the clusters/classes are predefined as following:

$$V_{kj} = \begin{cases} \alpha k = j \\ 0 k \neq j \end{cases}$$

Therefore, the centroid of the $j$-th cluster [with $j = 1 \ldots N_{sub}$; $N_{sub}$ is the total number of subjects (i.e., number of controls plus the patient)] is a vector of zeros except the $j$-th value which is set to $\alpha$ (with $\alpha \neq 0$). The real constant $\alpha$ is a “tuning” parameter that (1) adjusts the sensitivity of the method to outlier values, and (2) its sign determines whether to look for abnormally high ($\alpha > 0$) or low ($\alpha < 0$) outlier values; here $\alpha$ is fixed at $-0.5$, cf. Seghier et al. (2008).

We then quantify a similarity metric $D_{ij}$ at each voxel $i$ of the brain of the $j$-th subject:

$$D_{ij} = 1 - \tanh \left( \frac{N_{sub}}{\alpha^2} \cdot \text{cov}(P_i, V_j) \right)$$

$tanh$ is the hyperbolic tangent, $\text{cov}$ is the covariance between $P_i$ (a vector of tissue probability at voxel $i$ across all subjects) and $V_j$ (the centroid of the $j$-th cluster). The distance $D_{ij}$ lies in the range zero to two and will be small when $P_i$ covaries highly with $V_j$; this means the effects in voxel $i$ is driven or influenced by the $j$-th subject. The above formula can be rewritten as follows (cf. Seghier et al., 2007):

$$D_{ij} = 1 - \tanh \left( \frac{N_{sub} \cdot \bar{P}_{ij} - \bar{P}_j}{\alpha} \right)$$

Where $\bar{P}_{ij}$ is the mean of the tissue probability at voxel $i$ over all subjects excluding the $j$-th subject, and likewise $\bar{P}_j$ is the mean of the tissue probability at voxel $i$ over all subjects including the $j$-th subject. Accordingly, at a given voxel $i$, $D_{ij}$ is a measure of (1) how far subject $j$ is from the group mean, and (2) how the group mean is perturbed when subject $j$ is included. This observation is important and suggests that the FCP algorithm is formally similar to regression diagnostic methods that assess the extent to which a particular data point influences the model, by determining the change when that point is omitted (see a detailed discussion in Seghier et al., 2007).

The similarity metric $D_{ij}$ is then used to quantify the degree of membership $U_{ij}$ of voxel $i$ to cluster $j$ as follows:

$$U_{ij} = \frac{D_{ij}^4}{\sum_j D_{ij}^4}$$

Where the real constant $\lambda$ is a “defuzzification” parameter: when $\lambda$ tends to $-\infty$ the classification becomes crisp and $U_{ij}$ takes the value 0 or 1, but when $\lambda$ goes to 0 the classification is fuzzy ($U_{ij}$ is near to $1/N_{sub}$). We held $\lambda$ constant at $-4$ (cf. Seghier et al., 2007, 2008).

When $j$ indexes the patient, the set of $U_{ij}$ values comprise a GM (or WM) lesion image, which represents the degree of membership (i.e., abnormality) of voxels that have very low GM (or WM) probability in the patient relative to controls. Both GM and WM lesion images are then combined to obtain a single lesion image (as a fuzzy set) that codes the degree of abnormality at each voxel of the brain.

**Recursive ALI**

When ALI was tested with its default parameterization on patients with brain tumors, it identified some tumor tissue at the correct location, but not always with the correct extent. Specifically, in those cases, the segmentation step with the default extra prior failed to detect all the atypical tissue, and hence misclassified tumor voxels as either intact GM or WM (see Figure 2, top panel). The outlier detection algorithm (step “B” as detailed above) cannot recover the misclassified tumor voxels and thus we needed first to ensure accurate brain segmentation (step “A”) in the presence of tumors. This can be achieved by an optimization of the priors used to constrain tissue classification. One obvious possibility was to increase the number of iterations (i.e., number of segmentation runs in step “A”). This is because the prior of the extra class is optimized iteratively in the segmentation step of ALI (see Figure 1); hence more iterations could theoretically increase the extent of the abnormal tissue in the extra tissue class. However, this was not the case with some of the images from patients with tumors because the extra prior was empty (or including few voxels with very weak probabilities) after the first iteration, making any additional iterations to the segmentation step ineffective.

Our solution to this problem, in the segmentation step, was motivated by the fact that, in a Bayesian sense, posterior estimates of the tissue classes improve with the accuracy of the definition of the spatial priors (see for example Levy-Cooperman et al., 2008). As mentioned above, we first noticed that tumors were being identified at the correct location with ALI, but not always with the full extent. By using ALI in a recursive mode, we assumed that the output from a first run of ALI would serve as a good empirical definition of the prior for the extra class (See Figure 1, Figure 2).
red arrow). ALI could then be repeated with the same parameters but with the inclusion of a refined patient-specific prior for the extra class that corresponded to the fuzzy set lesion image obtained during the first run (see Figure 2, bottom panel). In other words, we suggest a “hierarchical” framework where the output from a first run of ALI serves as a prior in a second (or more) run of ALI. This extra prior is expected to grow nonlinearly with the number of iterations in the segmentation step in order to encompass the full extent of the lesion (as detailed above). Here we assume that two repetitions of ALI is near-optimal for our purposes. Conceptually, as reported in Cardoso et al. (2011), the aforementioned estimated priors are not proper priors from a mathematical standpoint, as they are derived from the data (i.e., generated by transforming the posteriors after the first run of ALI). However, they act as proper priors in the segmentation step of the second run of ALI. Below we show how running ALI in a recursive mode significantly improved the lesion definition in challenging cases while maintaining the automatic nature of the procedure.

VALIDATION

The outputs from our recursive ALI procedure were compared to the gold standard: manual segmentations of the same tumors. Practically, outline of the lesion was drawn on the outer borders of abnormally intense regions. The lesions were identified on a slice by slice basis on the T1-weighted images in the native ders of abnormally intense regions. The lesions were identified to the gold standard: manual segmentations of the same tumors. The outputs from our recursive ALI procedure were compared obtained during the first run (see the extra class that corresponded to the fuzzy set lesion image ters but with the inclusion of a refined patient-specific prior for the red arrow). ALI could then be repeated with the same parameters in a recursive mode significantly improved the lesion definition in challenging cases while maintaining the automatic nature of the procedure.

The Dice’s similarity index (Dice, 1945) was used to estimate the similarity between two binary images using the following formula:

\[
\text{Dice} = \frac{2TP}{2TP + FP + FN}
\]

Where the TP, FP, and FN represent the number of true positives (TP), false positives (FP) and false negatives (FN) in one approach relative to another. We calculated the Dice’s similarity index for (1) each binary lesion map of the automated procedure (i.e., fuzzy sets at a given U threshold) relative to the manually-segmented tumors of each observer, and (2) the manually-segmented tumors of one observer relative to that of the other. The higher the Dice’s index (ranging from 0 to 1) the higher the similarity between lesion definitions.

We also wanted to evaluate the sensitivity and specificity of our automated procedure in comparison to the manual segmentation of the same tumors. The sensitivity indicates the percentage of correctly classified voxels within the tumor (true positive rate), and the specificity represents the proportion of correctly classified non-tumor voxels (true negative rate). To estimate these values, we generated ROC curves (e.g., Metz, 1978) that encode the dependence of the true positive rate (sensitivity) on the false positive rate (one minus specificity) for different U thresholds. Perfect tumor segmentation would yield a point in the upper left corner of the ROC curve (i.e., no false negatives and no false positives). A completely random segmentation would result in a value along the diagonal line of the ROC curve. The area under the curve (AUC) (Hanley and McNeil, 1982) was assessed as a measure of the effectiveness of the recursive ALI method. Although the interpretation of the AUC might be problematic, particularly when comparing between models or classifiers (Lobo et al., 2008; Powers, 2012), it has nevertheless been widely used for assessing the accuracy of automated segmentation methods. Here, we provided both AUC values and Dice similarity indices in all patients as previously recommended (for a discussion see Zou et al., 2004). For completeness, the results were also reported in terms of sensitivity and specificity. Last but not least, AUC was obtained for both the standard and recursive ALI approaches, and a Wilcoxon signed rank test was used to assess whether the AUC values for our recursive approach were statistically larger than those obtained with the standard ALI.

RESULTS

At the global (visual inspection) level, our recursive ALI procedure was able to detect tumors at the right location and with the correct extent in all cases, except one (Patient 18). The boundaries identified for each tumor are shown in Figure 3 which demonstrates that the recursive ALI procedure successfully identified tumors of different types (low and high grade gliomas) and sizes, including tumors located in the insula (e.g., see Patients 1 and 16), close to the inter-hemispheric fissure (e.g., Patients 6 and 17) or near to the ventricles (e.g., Patient 11).
FIGURE 3 | Axial slices illustrating the lesion boundaries in each of the 18 patients as defined by the recursive ALI procedure detailed in Figures 1, 2. The tumor contour was obtained with a threshold of $U > 0.3$. L, Left; R, Right.

At the voxel level, the Dice’s similarity index between the two observers was on average 0.838 (±0.110). Figure 4A illustrates the similarity, according to the Dice index, between each binary tumor image obtained with our recursive ALI method compared to the manually-segmented tumors. For each patient, the Dice similarity index was generated multiple times across a range of different $U$ thresholds used in the conversion of the fuzzy image into the binary image. The plots in Figure 4A illustrate that similarity with the manual segmentation approach was highest (higher Dice indices) when the threshold on the degree of abnormality $U$ ranged from 0.1 to 0.3. The ROC curves for all patients showed that our method is highly sensitive and specific in comparison to manual segmentation (see Figure 4B). The AUC from the ROC analysis for the recursive ALI was 0.97 on average (range = 0.90–0.99), which suggests an excellent agreement with the manual segmentation. It was also significantly higher (Wilcoxon signed rank test: z-value: 3.7, $p < 0.001$) when compared with the standard ALI (mean $AUC = 0.92$; range = 0.84–0.98). See also Table 1 for a list of all values for each individual patient including the Dice similarity index (at a $U$ threshold of 0.2), the AUC (for the standard and recursive ALI), and both sensitivity and specificity (at a $U$ threshold of 0.2 for the recursive ALI).

DISCUSSION
The present study aimed to demonstrate that brain tumors can be automatically identified, in an operator-independent way, from single MR images, using a modified version of a previously described procedure (Seghier et al., 2008). Our results show that this new procedure, which we refer to as recursive ALI, was able to delineate all the tumors at the right location, with high accuracy in 17 out of 18 different patients, despite the heterogeneous range of tumors included in our sample. The success of the tumor identification was validated in comparison to manual segmentation which is the current gold standard approach. We also show how tumor identification, in some patients, improved with the new recursive-ALI approach compared to the standard ALI approach. The recursive ALI procedure that we propose here will be particularly useful for studies of lesion-behavior mappings with large samples of patients. It may also have potential uses for surgical or diagnostic purposes.

Our recursive ALI procedure was able to identify tumors at the correct location in all the patients irrespective of the type, size or location. We stress here that the output was not determined by the particular low T1 signal in the tumors (as compared to the signal in the WM, see Figure 3). Successful automated delineation is also possible even if the tumors exhibited hypersignal because the critical point is not how darker or brighter the tumors appear in MRI scans but whether the signal is abnormal in comparison to what is typically expected at that particular location. It is also important to keep in mind that the outlier detection algorithm operates on the segmented GM and WM tissue classes rather than the raw T1 signal. Additionally, our method identified the extent of the
FIGURE 4 | (Top Panel) Dice’s similarity index when the binary lesion from the recursive ALI procedure is compared to each observer’s manual segmentation (left = Observer 1; right = Observer 2), across a range of different U thresholds used to convert the fuzzy image into a binary image. The different colors represent different patients. (Bottom panel) ROC curves for different U thresholds. All the curves are close to the top-left corner (near to the manual segmentation). Patient 18 (with the most challenging tumor definition) is displayed in turquoise.

tumor in all but one case (Patient 18 who had a high grade tumor, see below). The AUC from the ROC analysis was significantly higher for recursive ALI than for standard ALI. Notably, recursive ALI showed the biggest improvements when tumor delineation was challenging for the standard procedure (Patients 6 and 16). Patient 6 had extensive oedema near the inter-hemispheric fissure. Patient 16 had a tumor with a signal distribution close to the normal GM and located in the insula (Figure 2) where high inter-subject variability in normal anatomy limits the effectiveness of the segmentation priors. In sum, identifying brain tumors with recursive ALI is equal to or better than that obtained with standard ALI. Therefore, it may also be appropriate for the identification of other types of lesions that are problematic in the standard ALI implementation.

When comparing recursive ALI to the manually segmented tumors, the Dice similarity index (at a U threshold of 0.2) and the AUC were on average 0.72 and 0.97, respectively (See Figure 4 and Table 1). Therefore, both indices suggest an excellent agreement between automated and manual segmentations performed by each observer for the same tumors. To interpret our results, it is important to keep in mind the inherent inter-operator variability when using manual tracing as the gold standard method. By comparing manual lesion segmentation between and within observers, previous studies have shown non-negligible intra- and inter-operator variability for lesion identification (Fiez et al., 2000; Kaus et al., 2001). For instance, in our data, Dice’s similarity index between the two independent observers was 0.838 on average, with higher variability across patients with diffuse
signal (e.g., Patients 3 and 18, Table 1). In the same way, by defining variability as the percentage of non-overlapping voxels between two manually segmented lesions (cf. Fiez et al., 2000), the inter-operator variability was on average 27% (SD = 15%) in our data. This figure is comparable to the one reported by Fiez et al. (2000) at 33% (SD = 7%) in inter-operator variability between two observers segmenting MRI images from 10 patients with left frontal lesions. This indicates that (1) the definition of true positives and negatives is contaminated by subjectivity in the observers’ manual segmentation. This would yield lower sensitivity and specificity for the recursive ALI method because, by definition, any alternative method cannot outperform the gold standard method and (2) our procedure would be a reasonable substitute for an observer, since the differences between the automated and the manual method are at a similar magnitude as the differences between observers, but with the advantage of being unbiased and replicable.

Recursive ALI may also detect secondary lesion effects such as atypical enlarged ventricles or widened sulci in the ipsi- and/or contra-lesional hemisphere. In our sample this was the case of Patient 2, in which enlarged ventricles were defined as lesion by the recursive ALI (see Figure 3). If these areas need to be omitted, then users can either (1) factor them out by including older healthy subjects who are likely to have visible aging-induced atrophy, or (2) adjust the threshold parameters when converting the fuzzy lesion image to the binary lesion image. The latter solution will only help, however, when the size of secondary effects are distant to and smaller than the primary damage. For example, in Patient 2, the ventricular dilations can easily be eliminated from the binary lesion image by increasing the cluster extent threshold. In the same way, the identification of tiny tumors (e.g., size <0.5 cm³) might be challenging because such small tumors express themselves at a spatial scale comparable to the expected normal variability in anatomy (few millimetres). As a general principle, for such challenging cases, we recommend all identified tumors be checked by eye.

Tumor identification in Patient 18 was the least successful (see Figure 4 and Table 1). This patient had a high grade tumor (glioblastoma) located in the temporo-parietal junction that infiltrated ventrally into the temporal pole and dorsally into the superior parietal lobule (Figure 5), resulting in an unclear border between intact and abnormal tissue. Consequently, the two observers disagreed on the exact borders of the tumor (see illustration in Figure 5), causing a relatively low Dice index (0.60). This situation is not an unusual occurrence in manual segmentation, even in studies with more than two observers (e.g., Kaus et al., 2001; Kubben et al., 2010). One possibility is to define the proper lesion as the set of voxels with the largest overlap over observers’ manual segmentation, a rationale equivalent to “the consensus lesion” suggested previously by Fiez et al. (2000). By taking the overlap between the two observers’ manual segmentation as the best approximation of Patient 18’s tumor, we found that the segmentation step misclassified some of the tumor as “intact” GM even with the inclusion of the patient-specific prior (see red contours in Figure 5). Overall, the difficulty delineating the exact tumor site in Patient 18, irrespective of whether tumor identification was conducted automatically or manually, highlights (1) the importance of visual inspection following automated tumor identification and (2) the limitation of relying on a single MRI image. The inclusion of other brain images (e.g., contrast-enhanced images, FLAIR, T2-weighted or DWI images) may be required in these cases (see below). Alternatively, we recommend (1) adjusting the probability threshold of the extra prior in the iterative segmentation step to avoid misclassification of abnormal voxels as either GM or WM, (2) adjusting the parameters of the fuzzy clustering algorithm for outlier detection, or (3) repeating the recursive procedure more than twice to refine the patient specific prior for the extra class. Note, however, that in the latter case, multiple repetitions are time consuming and might also yield misclassification of intact tissue into the extra prior. For a discussion on the effects of the number of iterations and other parameters on the lesion identification algorithm see Seghier et al. (2007, 2008). These adjustments may be particularly important for our automated tumor identification which is expected to be problematic when there is an infiltrative tumor with a complex texture and unclear boundaries (Patient 18) or when tissue misclassification occurs in areas with high anatomical variability between controls (Patient 16).

Although the ALI procedure identifies lesions at the individual patient level, it is critically dependent on accurate spatial normalization. This is because the segmentation priors are in standard MNI space and also because the comparison of GM and WM images in the patient and controls during outlier detection relies on accurate warping of lesioned and healthy brains to the same standard reference space. The challenges of spatial normalization in brain images with lesions are well-documented (Brett et al., 2001; Crinion et al., 2007; Andersen et al., 2010; Ripollès et al., 2012). Our proposed recursive ALI procedure may help...
to improve spatial normalization of lesional brains. Specifically, by optimizing an extra prior, at the individual patient level, the recursive ALI procedure may reduce tissue misclassification and thus may enhance the precision of the warping. This would ensure accurate spatial normalization, as implemented within the unified segmentation-normalization framework (Ashburner and Friston, 2005), yielding higher statistical power in anatomical and functional imaging studies of groups of patients with brain lesions.

The results of this study also offer another illustration of the flexibility of ALI when dealing with different types of brain damage. Indeed, modified versions of ALI have been successfully applied in clinical contexts, for instance in the identification of cerebral microbleeds from low-resolution T2*-weighted images (Seghier et al., 2011), the delineation of lesions from clinical CT scans (Chechlacz et al., 2012), and the detection of corticospinal tract damage in stroke patients (Kou et al., 2013). Its robustness on monospectral data makes ALI an attractive tool for other lesion types that can be fully characterized with a single imaging modality. This includes, for instance, the detection of white matter hyperintensities that typically appear on T2-weighted fluid attenuated inversion recovery (FLAIR) images (Wen and Sachdev, 2004), where automated identification (e.g., Wu et al., 2006a) of such white matter hyperintensities can be clinically useful in assessing risk factors of stroke and dementia (see meta-analysis in Debette and Markus, 2010).

Our point here is that the segmentation procedure used in ALI can be optimized for a given type of brain damage by explicitly incorporating any prior knowledge that can be provided by theoretical or other empirical data. For example, in the context of brain tumors, it is perfectly possible to define subject-specific priors that could be derived from other contrast images that are likely to be available in the clinical setting (e.g., contrast-enhanced images, FLAIR, T2-weighted or DWI images, see Prastawa et al., 2003), or even include more than one extra class to model the different types of the abnormal tissue (e.g., see Zacharakis et al., 2012). Indeed, the accuracy of ALI might be boosted by incorporating other contrast images because any additional information in those images may help the segmentation procedure to optimally differentiate between abnormal and normal brain tissue. When multi-contrast images are available, users may choose to run ALI (i.e., the segmentation step) in a multispectral mode, provided all contrast images are spatially co-registered and resliced. The effectiveness of multispectral segmentation in ALI for tumor identification warrants future investigations. An accurate delineation of the tumor can then be combined with pre-operative fMRI mapping of the same patients to support neurosurgical planning and enhance the efficiency and safety of the image-guided procedures (e.g., Tieleman et al., 2009; Kekhia et al., 2011).

Finally, although the output from our recursive ALI procedure is a fuzzy set indexing the degree of abnormality at each voxel of the brain, it is also possible to use the final segmented extra class as a proxy of the tumor that might be particularly useful for people interested in probabilistic estimates of tumors. We found very good correspondence between the segmented tissue in the extra class (i.e., the output of the segmentation during the second run of ALI) and the manual lesion segmentation of each patient (e.g., see Figure 2). The use of the fuzzy clustering algorithm, however, allows patient-specific tissue classes to be compared to a set of images from matched healthy controls (e.g., with similar demographics and/or acquisition protocols). Last but not least, we found that Dice indices were higher when the threshold on the degree of abnormality U ranged from 0.1 to 0.3 (Figure 4).

For similar datasets, we thus recommend the use of U thresholds within this range if users are interested in generating binary definitions of their patients’ tumors. This threshold needs to be increased (e.g., at 0.5) if a small number of controls were used in the outlier detection algorithm. As the case for any method that relies upon fuzzy or soft clustering techniques, the optimal threshold can be adjusted for a given dataset (see Shen et al., 2010 for a discussion).

To conclude, we have optimized our automated lesion identification procedure for brain tumors. This involved improving the tissue segmentation by including a data-driven recursive step that generated a patient-specific prior for the lesioned tissue. The rationale for developing this procedure was based on the need to identify tumors from monospectral MRI images when additional MRI contrast images are not available. Our findings demonstrate the success of the recursive ALI procedure by delineating different brain tumors with very good accuracy in a sample of 18 patients whose tumors differed in type, size and location. The recursive ALI procedure we introduce here therefore offers an objective and replicable method for brain tumor identification that will be relevant for lesion-behavior mapping. In addition, it may also be applied for the identification of other types of lesions that are problematic in the standard ALI implementation and may help to achieve accurate spatial normalization of brains with tumors for robust group analyses. The present paper also illustrates that ALI is a versatile tool that can be adapted for automated identification of any type of brain damage.

AUTHORS CONTRIBUTIONS
Ana Sanjuán, Cathy J. Price, and Tarek A. Yousry conceived and designed the experiment; Ana Sanjuán analyzed the data; Mohamed L. Seghier designed the toolbox used in analysis; Ana Sanjuán and Sharon Geva manually segmented the tumors; Laura Mancini, Goulven Josse, Alice Grogan, Adam K. Yamamoto, Alex P. Leff and Tarek A. Yousry contributed with patient selection and data acquisition; Ana Sanjuán, Cathy J. Price, and Mohamed L. Seghier wrote the paper; Ana Sanjuán, Cathy J. Price, Laura Mancini, Adam K. Yamamoto, Sharon Geva, Alex P. Leff, Tarek A. Yousry, and Mohamed L. Seghier revised the manuscript.

ACKNOWLEDGMENTS
Data collection was funded by the Wellcome Trust, the James S. McDonnell Foundation (as part of the brain network recovery group) and funding from the Comprehensive Biomedical Research Council. Data analysis was funded by the County Council Department of Education, Culture and Sports of the Valencian Community, Spain (grant number APOSTD/2012068) and the Universitat Jaume I, Spain (grant number P1-1B2012-38) awarded to the first author. A beta version of ALI is available.
Seghier, M. L., Friston, K. J., and Price, C. J. (2007). Detecting subject-specific activations using fuzzy clustering Neuroimage 36, 594–605. doi: 10.1016/j.neuroimage.2007.03.021

Seghier, M. L., Kolanko, M. A., Leff, A. P., Jäger, H. R., Gregoire, S. M., and Werring, D. J. (2011). Microbleed detection using automated segmentation (MIDAS): a new method applicable to standard clinical MR images. PLoS ONE 6:e17547. doi: 10.1371/journal.pone.0017547

Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., and Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage 41, 1253–1266. doi: 10.1016/j.neuroimage.2008.03.028

Shen, S., Szameitat, A. J., and Sterr, A. (2010). An improved lesion detection approach based on similarity measurement between fuzzy intensity segmentation and spatial probability maps. Magn. Reson. Imaging 28, 245–254. doi: 10.1016/j.mri.2009.06.007

Soltanian-Zadeh, H., Peck, D. J., Windham, J. P., and Mikkelsen, T. (1998). Brain tumor segmentation and characterization by pattern analysis of multi-spectral NMR images. NMR Biomed. 11, 201–208. doi: 10.1002/(SICI)1099-1492(199806/08)11:4/5<201::AID-NBM508>3.0.CO;2-6

Thesen, T., Quinn, B. T., Carlson, C., Devinsky, O., DuBois, J., McDonald, C. R., et al. (2011). Detection of epileptogenic cortical malformations with surface-based MRI morphometry. PLoS ONE 6:16430. doi: 10.1371/journal.pone.0016430

Tieleman, A., Deblaere, K., Van Roost, D., Van Damme, O., and Achten, E. (2009). Preoperative fMRI in tumour surgery. Eur. Radiol. 19, 2523–2534. doi: 10.1007/s00330-009-1429-z

Veloz, A., Orellana, A., VieIma, J., Salas, R., and Chabert, S. (2011). “Brain tumours: How can images and segmentation techniques help?” in Diagnostic Techniques and Surgical Management of Brain Tumors, ed A. L. Abujamra (InTech), 67–92. doi: 10.5772/22466. Available online at: http://www.intechopen.com/books/diagnostic-techniques-and-surgical-management-of-brain-tumors/brain-tumors-how-can-images-and-segmentation-techniques-help

Wen, W., and Sachdev, P. (2004). The topography of white matter hyperintensities on brain MRI in healthy 60- to 64-year-old individuals. Neuroimage 22, 144–154. doi: 10.1016/j.neuroimage.2003.12.027

Wu, M., Rosano, C., Butters, M., Whyte, E., Nable, M., Crooks, R., et al. (2006a). A fully automated method for quantifying and localizing white matter hyperintensities on MR images. Psychiatry Res. 148, 133–142. doi: 10.1016/j.psychres.2006.09.003

Xu, T., and Mandal, M. (2012). Automatic brain tumor extraction from T1-Weighted coronal mri using fast bounding box and dynamic snake. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2012, 444–447.

Zacharakis, E., Erus, G., Bezerianos, A., and Davatzikos, C., (2012). “Fuzzy Multi-channel clustering with individualized spatial priors for segmenting brain lesions and infarcts,” in Artificial Intelligence Applications and Innovations. In IFIP Advances in Information and Communication Technology, Vol. 382, (Halkidiki: Springer Berlin Heidelberg), 76–85. doi: 10.1007/978-3-642-33412-2_8

Zou, K. H., Wells, W. M. 3rd, Kikinis, R., and Warfield, S. K. (2004). Three validation metrics for automated probabilistic image segmentation of brain tumours. Stat. Med. 23, 1259–1282. doi: 10.1002/sim.1723

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 August 2013; paper pending published: 19 September 2013; accepted: 27 November 2013; published online: 17 December 2013.

Citation: Sanjuán A, Price CJ, Mancini L, Josse G, Grogan A, Yamamoto AK, Geva S, Leff AP, Yousry TA and Seghier ML (2013) Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors. Front. Neurosci. 7:241. doi: 10.3389/fnins.2013.00241

This article was submitted to Brain Imaging Methods, a section of the journal Frontiers in Neuroscience.

Copyright © 2013 Sanjuán, Price, Mancini, Josse, Grogan, Yamamoto, Geva, Leff, Yousry and Seghier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.
