Contents lists available at ScienceDirect

Data in Brief

Data Article

Longitudinal multiple sclerosis lesion segmentation data resource

Aaron Carass [a,b,*], Snehashis Roy [c], Amod Jog [b],
Jennifer L. Cuzzocreo [d], Elizabeth Magrath [c], Adrian Gherman [e],
Julia Button [d], James Nguyen [d], Pierre-Louis Bazin [f],
Peter A. Calabresi [d], Ciprian M. Crainiceanu [e],
Lotta M. Ellingsen [a,g], Daniel S. Reich [d,h], Jerry L. Prince [a,b],
Dzung L. Pham [c]

[a] Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD 21218, USA
[b] Department of Computer Science, The Johns Hopkins University, Baltimore, MD 21218, USA
[c] CNRM, The Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD 20892, USA
[d] Department of Radiology, The Johns Hopkins School of Medicine, Baltimore, MD 21287, USA
[e] Department of Biostatistics, The Johns Hopkins University, Baltimore, MD 21205, USA
[f] Department of Neurophysics, Max Planck Institute, 04103 Leipzig, Germany
[g] Department of Electrical and Computer Engineering, University of Iceland, 107 Reykjavík, Iceland
[h] Translational Neuroradiology Unit, National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA

ARTICLE INFO

ABSTRACT

The data presented in this article is related to the research article entitled "Longitudinal multiple sclerosis lesion segmentation: Resource and challenge" (Carass et al., 2017) [1]. In conjunction with the 2015 International Symposium on Biomedical Imaging, we organized a longitudinal multiple sclerosis (MS) lesion segmentation challenge providing training and test data to registered participants. The training data consists of five subjects with a mean of 4.4 ($+0.55$) time-points, and test data of fourteen subjects with a mean of 4.4 ($+0.67$) time-points. All 82 data sets had the white matter lesions associated with multiple sclerosis delineated by two human expert raters. The training data including multi-modal scans and manually delineated lesion masks is available for download.[1] In addition, the

testing data is also being made available in conjunction with a website for evaluating the automated analysis of the testing data.

Specifications Table

Subject Area	Neurology
More Specific Subject Area	Neuroimaging
Type of Data	Magnetic Resonance Images Specifically:
	T_1-w MPRAGE
	T_2-w & PD-w DSE
	T_2-w FLAIR
Data Format	Raw and Processed
Experimental Factors	None
Data Source Location	The Johns Hopkins Hospital, Baltimore, MD 21287
Data Accessibility	Public download

Value of the data

- This is currently the largest available public database of manually delineated MS lesions.
- All 82 data sets have been manually delineated by two raters.
- A unique multi-modal data set of MS lesion progression covering multiple time-points.
- A public evaluation website allows for the comparison of automated algorithms on the testing data.

1. Data

The data presented in this article is related to the research article entitled "Longitudinal multiple sclerosis lesion segmentation: Resource and challenge" [1]. The data consists of magnetic resonance (MR) images (MRI) divided into two cohorts: 1) Training Set; and 2) Test Set. The Training Set consists of five subjects, four of which had four time-points, while the fifth subject had five time-points. The Test Set includes fourteen subjects, ten of which had four time-points, three had five time-points, and one had six time-points. Two consecutive time-points are separated by approximately one year for all subjects. Table 1 includes a demo-graphic breakdown for the training and test data sets. The data does not supply the multiple sclerosis (MS) subtype of the subjects for either the training or the test data. The data is available for download from the Challenge Evaluation Website: http://smart-stats-tools.org/lesion-challenge-2015.

2. Methods

Each scan was imaged and preprocessed in the same manner, with data acquired on a 3.0 T MRI scanner (Philips Medical Systems, Best, The Netherlands) using the following sequences: a T_1-weighted (T_1-w) magnetization prepared rapid gradient echo (MPRAGE) with TR = 10.3 ms, TE = 6 ms, flip angle = 8°, & $0.82 \times 0.82 \times 1.17$ mm^3 voxel size; a double spin echo (DSE) which produces the proton density weighted (PD-w) and T_2-weighted (T_2-w) images with TR = 4177 ms, TE$_1$ = 12.31 ms, TE$_2$ = 80 ms, & $0.82 \times 0.82 \times 2.2$ mm^3 voxel size; and a T_2-w fluid attenuated inversion recovery (FLAIR) with

Table 1
Demographic details for the training data and both test data sets. The top line is the information of the entire data set, while subsequent lines within a section are specific to the patient diagnoses. The codes are RR for relapsing remitting MS, PP for primary progressive MS, and SP for secondary progressive MS. N (M/F) denotes the number of patients and the male/female ratio, respectively. Time-points is the mean (and standard deviation) of the number of time-points. Age is the mean age (and standard deviation), in years, at baseline. Follow-up is the mean (and standard deviation), in years, of the time between follow-up scans.

Data set	N (M/F)	Time-points mean (SD)	Age mean (SD)	Follow-up mean (SD)
Training	5 (1/4)	4.4 ($\pm$0.55)	43.5 ($\pm$10.3)	1.0 ($\pm$0.13)
RR	4 (1/3)	4.5 ($\pm$0.50)	40.0 ($\pm$7.6)	1.0 ($\pm$0.14)
PP	1 (0/1)	4.0	57.9	1.0 ($\pm$0.04)
Test	14 (3/11)	4.4 ($\pm$0.63)	39.3 ($\pm$8.9)	1.0 ($\pm$0.23)
RR	12 (3/9)	4.4 ($\pm$0.67)	39.2 ($\pm$9.6)	1.0 ($\pm$0.25)
PP	1 (0/1)	4.0	39.0	1.0 ($\pm$0.04)
SP	1 (0/1)	4.0	41.7	1.0 ($\pm$0.05)
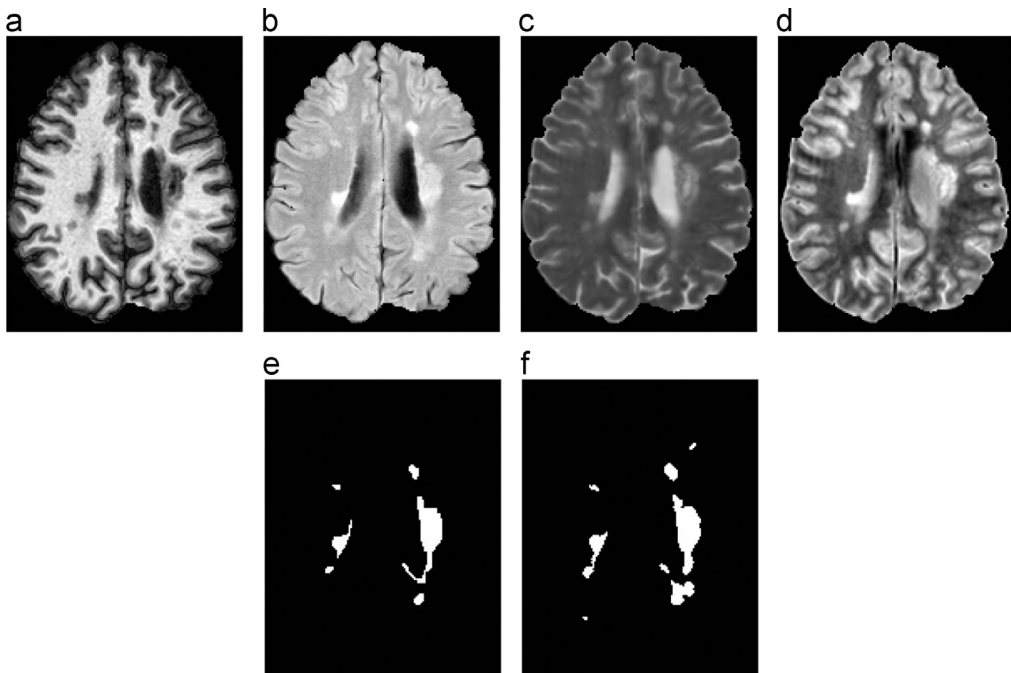
Fig. 1. Shown are the preprocessed (a) MPRAGE, (b) FLAIR, (c) T_2-w, and (d) PD-w images for a single time-point from one of the provided Training Set subjects. The corresponding manual delineations by our two raters are shown in (e) for Rater #1 and (f) for Rater #2.

TI=835 ms, TE=68 ms, & $0.82 \times 0.82 \times 2.2$ mm^3 voxel size. The imaging protocols were approved by the local institutional review board. Each subject underwent the following preprocessing: the baseline (first time-point) MPRAGE was inhomogeneity-corrected using N4 [2], skull-stripped [3,4], dura stripped [5], followed by a second N4 inhomogeneity correction, and rigid registration to a 1 mm isotropic MNI template [6]. We have found that running N4 a second time after skull and dura stripping is more effective (relative to a single correction) at reducing any inhomogeneity within the images (see Fig. 1 for an example image set after preprocessing). Once the baseline MPRAGE is in MNI space, it is used as a target for the remaining images. The remaining images include the baseline T_2-w, PD-w, and FLAIR, as well as the scans from each of the follow-up time-points. These images are N4 corrected and

then rigidly registered to the 1 mm isotropic baseline MPRAGE in MNI space. Our registration steps are inverse consistent and thus any registration based biases are avoided [7] The skull & dura stripped mask from the baseline MPRAGE is applied to all the subsequent images, which are then N4 corrected again.

For each time-point of every subject's scans in the Training Set and Test Set, the following data are provided: the original scan images consisting of T_1-w MPRAGE, T_2-w, PD-w, and FLAIR, as well as the preprocessed images (in MNI space) for each of the scan modalities. The Training Set also included manual delineations by two experts identifying and segmenting the white matter lesions on the MR images.

To facilitate the dissemination of the data and promote the sharing of results we have created a website.[2] Visitors to the site can see a list of the Top 25 submitted results. Groups interested in running their methods on the data need only register for an account, download the data, and upload their results. The uploader of the results will receive an e-mail within ten minutes detailing the results on a per subject and per time-point basis. The report includes the following computed metrics: Dice, Jaccard, positive predictive value (PPV), true positive rate (TPR), lesion false positive rate (LFPR), lesion true positive rate (LTPR), absolute volume difference (AVD), average symmetric surface distance (ASSD), algorithm and manual lesion volume. The website also provides an overall score; for a given algorithm A this is computed as follows,

$$\frac{1}{|\mathcal{R}|} \frac{1}{|\mathcal{S}|} \left(\sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{S}} \frac{\text{Dice } (\mathcal{M}_r, \mathcal{M}_A)}{8} + \frac{\text{PPV}(\mathcal{M}_r, \mathcal{M}_A)}{8} + \frac{1 - \text{LFPR}(\mathcal{M}_r, \mathcal{M}_A)}{4} + \frac{\text{LTPR}(\mathcal{M}_r, \mathcal{M}_A)}{4} + \frac{\text{Corr}(\mathcal{M}_r, \mathcal{M}_A)}{4} \right),$$

where $\mathcal{S}$ is the set of all subjects, $\mathcal{R}$ is the set of all raters, $\mathcal{M}_r$ is the lesion mask from rater r, $\mathcal{M}_A$ is the lesion mask for the algorithm A, and Corr is the Pearson's correlation coefficient of the volumes. This is then linearly normalized by the inter-rater scores between each rater such that the lower inter-rater score has an overall rating of 90. This was designed to mimic the scoring of the 2008 MICCAI MS Lesion challenge [8].

Acknowledgments

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2017.04.004.

References

[1] A. Carass, S. Roy, A. Jog, J.L. Cuzzocreo, E. Magrath, A. Gherman, J. But- ton, J. Nguyen, F. Prados, C.H. Sudre, N. Cawley, M. J. Cardoso, O. Cic- carelli, C.A.M. Wheeler-Kingshott, L. Catanese, S. Ourselin, H. Deshpande, P. Maurel, O. Commowick, C. Barillot, X. Tomas-Fernandez, S.K. Warfield Vaidya, R. Muthuganapathy, A. Chunduru, G. Krishnamurthi, A. Jesson Arbel, H. Handels, O. Maier, L.O. Iheme, D. Unay, S. Jain, D.M. Sima, M. Ghafoorian, D. Smeets, B. Platel, A. Birenbaum, H. Greenspan, P.A. Cal- abresi, P.-L. Bazin, C. Crainiceanu, L.M. Ellingsen, D.S. Reich, J.L. Prince, D.L. Pham, Longitudinal multiple sclerosis lesion segmentation: resource & challenge, NeuroImage (2017, 148(c), 2017) 77–102.
[2] N.J. Tustison, B.B. Avants, P.A. Cook, Y. Zheng, A. Egan, P.A. Yushkevich, J.C. Gee, N4ITK: improved N3 bias correction, IEEE Trans. Med. Imag. 29 (6) (2010) 1310–1320.
[3] A. Carass, M.B. Wheeler, J. Cuzzocreo, P.-L. Bazin, S.S. Bassett, J.L. Prince, A joint registration and segmentation approach to skull stripping, in: Proceedings of the 4th International Symposium on Biomedical Imaging, ISBI 2007, IEEE, 2007, pp. 656–659.

---

[2] The data and evaluation website is: http://smart-stats-tools.org/lesion-challenge-2015.

[4] A. Carass, J. Cuzzocreo, M.B. Wheeler, P.-L. Bazin, S.M. Resnick, J.L. Prince., Simple paradigm for extra-cerebral tissue removal: algorithm and analysis, NeuroImage 56 (4) (2010) 1982–1992.

[5] N. Shiee, P.-L. Bazin, J.L. Cuzzocreo, C. Ye, B. Kishore, A. Carass, P.A. Calabresi, D.S. Reich, J.L. Prince, D.L. Pham, Reconstruction of the human cerebral cortex robust to white matter lesions: method and validation, Hum. Brain Mapp. 35 (7) (2014) 3385–3401.

[6] V.S. Fonov, A.C. Evans, R.C. McKinstry, C.R. Almli, D.L. Collins, Unbiased nonlinear average age-appropriate brain templates from birth to adulthood, NeuroImage 47 (S1) (2009) S102.

[7] M. Reuter, B. Fischl, Avoiding asymmetry-induced bias in longitudinal. Image processing, NeuroImage 57 (1) (2011) 19–21.

[8] M. Styner, J. Lee, B. Chin, M.S. Chin, O. Commowick, H.-H. Tran, S. Markovic-Plese, V. Jewells, S. Warfield, 3D segmentation in the clinic: a grand challenge II: MS lesion segmentation, in: Proceedings of the 11th International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI 2008, 2008, pp. 1–6.