Modeling the Variability in Brain Morphology and Lesion Distribution in Multiple Sclerosis by Deep Learning

Tom Brosch[1,4], Youngjin Yoo[1,4], David K.B. Li[2,4], Anthony Traboulsee[3,4], and Roger Tam[2,4]

[1] Department of Electrical and Computer Engineering, UBC
[2] Department of Radiology, UBC
[3] Division of Neurology, UBC
[4] MS/MRI Research Group, University of British Columbia, Vancouver, Canada

Abstract. Changes in brain morphology and white matter lesions are two hallmarks of multiple sclerosis (MS) pathology, but their variability beyond volumetrics is poorly characterized. To further our understanding of complex MS pathology, we aim to build a statistical model of brain images that can automatically discover spatial patterns of variability in brain morphology and lesion distribution. We propose building such a model using a deep belief network (DBN), a layered network whose parameters can be learned from training images. In contrast to other manifold learning algorithms, the DBN approach does not require a prebuilt proximity graph, which is particularly advantageous for modeling lesions, because their sparse and random nature makes defining a suitable distance measure between lesion images challenging. Our model consists of a morphology DBN, a lesion DBN, and a joint DBN that models concurring morphological and lesion patterns. Our results show that this model can automatically discover the classic patterns of MS pathology, as well as more subtle ones, and that the parameters computed have strong relationships to MS clinical scores.

Keywords: Population modeling, multiple sclerosis, T2 lesion, machine learning, brain imaging, MRI, deep learning, deep belief networks.

1 Introduction

Multiple sclerosis (MS) is an inflammatory and degenerative disease of the central nervous system with pathology that can be observed in vivo by magnetic resonance imaging (MRI). MS is characterized by the formation of lesions, primarily visible in white matter on conventional MRI, and the death of nervous tissue leading to global and regional atrophy. A number of imaging biomarkers, such as lesion volume and whole brain volume, have established their importance for the study of MS. However, MS is a complex disease whose pathological variability extends well beyond what can be captured by global and local volumetric measures. Methods based on statistics of diffeomorphisms have been used to discover patterns of regional atrophy [1], but they are not designed to directly

P. Golland et al. (Eds.): MICCAI 2014, Part II, LNCS 8674, pp. 463–470, 2014.

model morphological variability. It would be highly desirable to have a method that can automatically discover potentially important patterns of variability in brain morphology and lesion distribution, which would advance our understanding of the complex pathology of MS. In addition, the joint modeling of brain morphology and lesion distribution would further our knowledge of how these two key pathological features interact. However, this type of modeling is very challenging due to the high dimensionality of the data. In recent years, there has been an increased interest in biomarker discovery using manifold learning to form high-level representations of medical images [2,3,4]. Manifold learning is motivated by the assumption that the space of brain images can be modeled to some approximation by a low-dimensional manifold. Various methods for manifold learning have been proposed (e.g., locally linear embedding (LLE), Laplacian eigenmaps (LEM), Isomaps) [5], with Isomaps and LEM being the most popular for medical image analysis. Most such methods require a prebuilt proximity graph based on a selected distance measure, the choice of which can be crucial for manifold learning [6]. Aljabar et al. proposed a method [3] for the joint modeling of multiple image features of neonatal brains. First, separate manifolds based on features from geometric surface models, non-rigid deformations, and image intensities are learned, then a second dimensionality reduction step is used to combine the individual manifold parameters.

We present a new method for modeling the variability in the morphology and lesion distribution of a large set of MRIs of MS patients. Our method is built using a deep belief network (DBN) [7], a layered network whose parameters can be learned from training images. An advantage of DBNs over other manifold learning methods is that it does not require a prebuilt proximity graph, which is particularly beneficial for modeling lesions, because the spareness and randomness of MS lesions make defining a suitable distance measure challenging and potentially biasing. Instead, the DBN approach assumes that a particular lesion configuration is a sample from an unknown distribution of lesion configurations that can be parameterized by a relatively small set of lesion distribution parameters. We model morphological and lesion variability with two individual DBNs, then form a joint model by replacing the individual top network layers with a new layer that joins both DBNs, similarly to the work on the joint modeling of auditory and visual signals for speech recognition [8]. Our results show that this model can automatically discover the classic patterns of MS pathology, as well as more subtle ones, and that the distribution parameters computed are found to have strong relationships to MS clinical scores.

2 Methods

Our proposed model for pattern discovery consists of three main components (a) a model that aims to find patterns of morphological changes in deformation fields, (b) a model that aims to find patterns in the spatial distribution of lesions, and (c) a joint model that aims to find concurring deformation and lesion distribution patterns. The morphology model is learned from a set of displacement

fields that are calculated via non-rigid registration from a set of T1-weighted (T1w) brain MRIs $D \subset I$, $I = \{I_n \mid I_n \colon \Omega \mapsto \mathbb{R}\}$, $\Omega \subset \mathbb{R}^3$ and the ICBM 152 nonlinear atlas template image [9]. First, all images of the training set are aligned to MNI space by a 9 degree-of-freedom registration using FLIRT [10]. Then for each image $I_n \in D$, the displacement field u_n, $u \colon \Omega \mapsto \mathbb{R}^3$, that describes the non-rigid transformation from template coordinates to image coordinates is calculated using FNIRT [11], where the displacement field u_n is represented by a $40 \times 48 \times 22$ grid of 3D displacement vectors. We assume that the displacement fields u_n are samples from an unknown distribution $p(u \mid D_1, \dots)$ that can be parameterized by far fewer parameters than the dimensionality of the fields themselves. In practice, the user typically selects the number of parameters to represent the data being explored. The task of finding patterns is to discover the underlying probability density function $p(u \mid D_1, \dots)$, where the parameters $(D_1, \dots)^T$ represent the patterns of variability of the displacement field distribution. This allows us to compare the morphology of two brains at a very high level in terms of the distribution parameters of their displacement fields u_1 and u_2 given by $\mathbb{E}[D_1, \dots \mid u_1]$ and $\mathbb{E}[D_1, \dots \mid u_2]$. Furthermore, we can visualize the modes of morphological variability of MS brain images, by sampling the space of displacement fields spanned by $(D_1, \dots)^T$ by calculating the expectation $\mathbb{E}[u \mid D_1, \dots]$ for a range of values for $(D_1, \dots)^T$.

The probability density function $p(u)$ is modeled by a deep belief network (DBN) [7], a generative probabilistic graphical model consisting of one layer of observed random variables (visible units) and multiple layers of latent random variables (hidden units). In a DBN, each pair of adjacent layers of random variables form a restricted Boltzmann machine (RBM). The first RBM receives the displacement fields of a training set as input and reduces the dimensionality of each field by discovering patterns of similarity that are common within groups of displacement fields. Each subsequent RBM receives the hidden unit activations of the previous RBM as input, thus learning successively more complex and abstract patterns from the training data. In particular, we use a DBN with three strided convolutional RBMs[1] (sconvRBMs) and two dense RBMs [12] with 16 and 2 hidden units, respectively. In the following, we will briefly review the sconvRBM, as this model is less often described in the literature, followed by how the visible and hidden units of the DBN relate to displacement fields and displacement field distribution parameters. A sconvRBM is a type of RBM in which the probabilistic relationships between the visible and hidden units are expressed in terms of strided convolutions, a type of convolution that shifts the filter kernel as a sliding window with a step size or stride $s > 1$. Due to the much smaller number of trainable parameters compared to dense RBMs, sconvRBMs are best suited for learning low- to mid-level features from very high-dimensional data. Compared to other more commonly used convolution-based RBMs [13], an advantage of sconvRBMs is that inference is invertible, which allows the reconstruction of the visible units from the hidden unit activations. In our application,

[1] The three sconvRBMs have stride sizes of $2 \times 2 \times 1$, $2 \times 2 \times 2$, $1 \times 1 \times 1$, filter sizes of $10 \times 10 \times 7$, $10 \times 10 \times 10$, $3 \times 5 \times 3$, and 32, 64, 32 filters, respectively.

this would allow for the reconstruction of deformation fields from distribution parameters. The complete morphology DBN can be trained layer-by-layer by training each RBM individually using contrastive divergence [7]. Once the parameters of the DBN have been learned from training data, we can use the model for inference. Let $\boldsymbol{v}_{\mathrm{d},1}, \ldots, \boldsymbol{v}_{\mathrm{d},5}$, $\boldsymbol{h}_{\mathrm{d},1}, \ldots, \boldsymbol{h}_{\mathrm{d},5}$, and $\boldsymbol{\theta}_{\mathrm{d},1}, \ldots, \boldsymbol{\theta}_{\mathrm{d},5}$ denote the visible units, hidden units, and parameters, respectively, of each RBM of the morphology DBN. Then, for a given displacement field u_n, we can calculate the parameters $(D_1, \ldots)^T$ of $u \sim p(u \mid D_1, \ldots)$ with

$$(D_1, \ldots)^T = \mathbb{E}[D_1, \ldots \mid u_n] = \mathbb{E}[\boldsymbol{h} \mid \boldsymbol{v}_{\mathrm{d},5}, \boldsymbol{\theta}_{\mathrm{d},5}] \tag{1}$$

$$\boldsymbol{v}_{\mathrm{d},i+1} = \mathbb{E}[\boldsymbol{h} \mid \boldsymbol{v}_{\mathrm{d},i}, \boldsymbol{\theta}_{\mathrm{d},i}] \tag{2}$$

where $i \in [1, 4]$ and $\boldsymbol{v}_{\mathrm{d},1} = u_n$. Inversely, the mean displacement field $\bar{u}$ given the distribution parameters can be calculated by

$$\bar{u} = \mathbb{E}[u \mid D_1, \ldots] = \mathbb{E}[\boldsymbol{v} \mid \boldsymbol{h}_{\mathrm{d},1}, \boldsymbol{\theta}_{\mathrm{d},1}] \tag{3}$$

$$\boldsymbol{h}_{\mathrm{d},i} = \mathbb{E}[\boldsymbol{v} \mid \boldsymbol{h}_{\mathrm{d},i+1}, \boldsymbol{\theta}_{\mathrm{d},i+1}] \tag{4}$$

where $i \in [1, 4]$ and $\boldsymbol{h}_{\mathrm{d},5} = (D_1, \ldots)^T$.

The input into our lesion model is a set of 3D binary lesion masks $l_n \in I$, which have been created from T2-weighted (T2w) and PD-weighted (PDw) MRIs by experts using a semi-automatic method. All lesion masks are spatially aligned to MNI space using the transformations as calculated for the corresponding T1w images. Analogous to the morphology model, we assume that lesion masks l_n are samples from an unknown distribution $l_n \sim p(l \mid L_1, \ldots)$ that can be parameterized by only relatively few parameters $(L_1, \ldots)^T$. The task of finding lesion patterns is to discover the underlying probability density function $p(l \mid L_1, \ldots)$, where the parameters $(L_1, \ldots)^T$ represent patterns of variability of MS lesions. Similar to the morphology DBN, the lesion DBN consists of three sconvRBMs[2] and two dense RBMs with 16 and 2 hidden units, respectively. However, we modified the energy function of the sconvRBMs to account for the sparse activation of MS lesion masks. Large black regions without local structure can lead to random activations of the hidden units and consequently the learning of random filters. To overcome this problem, we propose to incorporate a region of interest (ROI) term into the energy equation of the sconvRBM, which allows constraining the filter learning process to a given ROI. This can be achieved by element-wise multiplication of the visible and hidden units with a binary mask, which sets the visible and hidden units outside of the ROI to zero, thereby removing their contribution to the energy of the model. The ROI was chosen to include all white matter lesions from the training set. Similarly to the morphology model, for a trained lesion DBN and a given lesion mask l_n, we can calculate the parameters $(L_1, \ldots)^T$ of $l_n \sim p(l \mid L_1, \ldots)$ in the same manner as in (1) and (2). Likewise, the mean lesion configuration $\bar{l}$ given the distribution parameters $(L_1, \ldots)^T$ can be calculated in the same manner as in (3) and (4).

[2] The three sconvRBMs have stride sizes of $4 \times 4 \times 2$, $2 \times 2 \times 2$, $2 \times 2 \times 2$, filter sizes of $20 \times 20 \times 10$, $14 \times 14 \times 10$, $10 \times 14 \times 6$, and 32, 64, 64 filters, respectively.

To discover concurring patterns of morphology and lesion distribution, we combine the morphology DBN and the lesion DBN to form the joint DBN, which defines the joint distribution $p(u, l \mid J_1, \dots)$. The joint DBN consists of two pathways, each corresponding to the first 4 layers of the morphology and lesion DBNs, respectively, and a 5th RBM layer with 4 hidden units, which replaces the 5th layer of the individual DBNs and combines the hidden unit activations of the 4th layer RBMs. That is, the 5th RBM defines the joint probability $p(\boldsymbol{v}_\mathrm{j}, \boldsymbol{h}_\mathrm{j} \mid \boldsymbol{\theta}_\mathrm{j})$, where $\boldsymbol{v}_\mathrm{j} = (\boldsymbol{h}_{\mathrm{d},4}^T, \boldsymbol{h}_{\mathrm{l},4}^T)^T$ and $\boldsymbol{h}_\mathrm{j} = (J_1, \dots)^T$ are the modes of variability of morphological and lesion distribution changes.

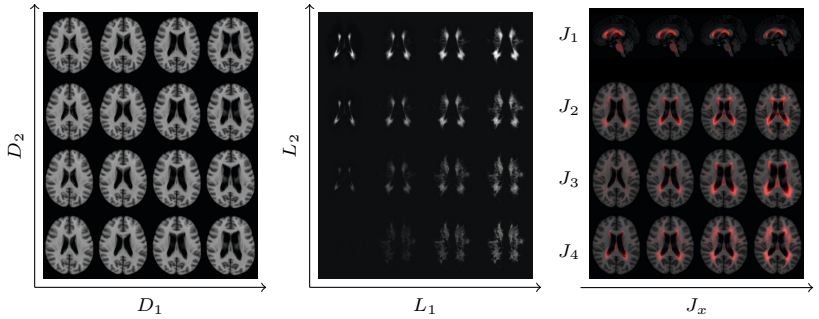
3 Experiments and Results

The proposed method was evaluated on a dataset from an MS clinical trial of 474 secondary progressive MS patients. For each subject, the dataset contains one T1w, T2w, and PDw MRI with a resolution of $256 \times 256 \times 50$ voxels and a voxel size of $0.937\,\mathrm{mm} \times 0.937\,\mathrm{mm} \times 3.000\,\mathrm{mm}$. The main preprocessing steps included rigid registration, brain extraction, intensity normalization, and background cropping. We then trained the 3 DBN models as described in Sect. 2.

The invertibility of our model allows the reconstruction of images from the distribution parameters to visualize the discovered patterns of variability. Figure 1(a) shows axial slices from volumes generated from the 2-parameter morphology model $p(u \mid D_1, D_2)$. To generate these images, we calculated the mean displacement fields for varying values of D_1 and D_2 to span the range of variability represented by the training set and applied the inverse deformations to the ICBM 152 template image. The most apparent morphological variability captured by the morphology model is ventricular enlargement for D_1 and overall brain size for D_2. Figure 1(b) shows axial slices from the mean lesion configurations $\mathbb{E}[l \mid L_1, L_2]$ for varying lesion distribution parameters. An increase of L_2 visually correlates with an increase of lesions specifically around the ventricles, whereas an increase of L_1 visually correlates with an increase of lesions in the entire white matter.

To visualize concurring patterns of morphology and lesion distribution, we sampled images from the joint model $p(u, l \mid J_1, \dots, J_4)$ as shown in Fig. 1(c). The images are deformed template images with superimposed lesion masks. For each row, we varied only one distribution parameter and set the remaining parameters to their mean values. Of the 4 parameters, J_3 visually corresponds most closely to the "classic" progression of MS pathology, with an enlargement of the ventricles paired with an increased periventricular lesion load. The parameters J_2 and J_4 also reveal simultaneous morphological and lesion variations that are visible on the chosen axial slice. For J_1, a lesion pattern is not obvious unless the images are viewed sagittally, which reveals changes in lesion load in the pons.

To evaluate the potential of the distribution parameters to reveal clinically relevant information, we have calculated the Pearson correlation r of the clinical MS



(a) Morphology manifold (b) Lesion manifold (c) Joint manifold

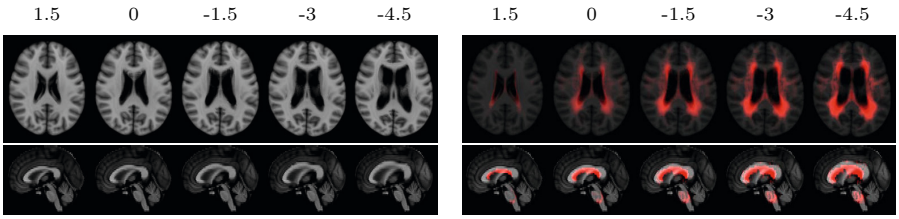
Fig. 1. Slices from generated volumes from the (a) morphology, (b) lesion, and (c) joint model. The morphology model captures ventricular enlargement (D_1) and decrease in brain size (D_2) as the main modes of variation. For the lesion model, L_1 captures an increase in lesion load throughout the WM, while L_2 captures primarily periventricular lesion load variations. The parameters of the joint model capture combinations of the variability found in the individual models.

Functional Composite (MSFC) [14] and its components (Timed 25-Foot Walk, T25W; 9-Hole Peg Test, 9-HPT; Paced Auditory Serial Addition Test, PASAT) with the distribution parameters and two established MS imaging biomarkers (normalized brain volume, nBV, as calculated by SIENAX [15] and lesion load, LL). The results of the correlation tests are summarized in Table 1. In the individual models, all parameters correlate significantly with 9-HPT, PASAT, and MSFC, except for D_2 with PASAT. The morphology parameter D_1 correlates more strongly with these scores than nBV, as does the lesion parameter L_2 than LL. For T25W, D_1 shows a modest but significant correlation while nBV does not. In the joint model, all parameters correlate significantly with 9-HPT, PASAT, and MSFC, with J_3 being particularly strong. The parameter J_3 shows stronger correlations than nBV or LL for all clinical scores, including a modest significant correlation to T25W, which is not shown by nBV nor LL. The significant correlation between J_1 and T25W is particularly interesting because the lesion changes modeled by J_1 occur in the pons, which is known to be of clinical significance for mobility. Overall, the learned parameters correlate better than the established imaging biomarkers.

Another benefit of our model is the ability to visualize the progression of a "mean" secondary progressive MS brain along a range of MSFC scores. To demonstrate, we trained 4 independent linear models to predict the distribution parameters $J_1, \ldots, J_4$ given the MSFC ($J_i = a_i + b_i\text{MSFC}$). Figure 2 shows the axial (top row) and mid-sagittal (bottom row) slices of generated images representing the range of MSFC scores from 1.5 to -4.5. Consistent with previous results, a decrease in MSFC visually correlates with an increase in the size of the ventricles, an increase in periventricular lesions, and an increase in lesions in the pons region.

Table 1. Pearson correlations r of clinical scores with distribution parameters of the morphology model (D_1, D_2), lesion model (L_1, L_2), joint model (J_1, J_2, J_3, J_4), normalized brain volume (nBV), and lesion load (LL). The level of statistical significance is indicated by the number of asterisks (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

		T25W	9-HPT	PASAT	MSFC
Individual models	D_1	**−0.129****	−0.215***	−0.282***	−0.315***
	D_2	0.087	0.116*	0.089	0.139**
	L_1	−0.058	−0.231***	−0.392***	−0.367***
	L_2	−0.091	**−0.354*****	**−0.427*****	**−0.464*****
Joint model	J_1	0.107*	0.286***	0.336***	0.379***
	J_2	−0.038	−0.210***	−0.227***	−0.256***
	J_3	**−0.118***	**−0.369*****	**−0.453*****	**−0.494*****
	J_4	−0.049	−0.206***	−0.383***	−0.346***
Imaging biomarkers	nBV	0.053	0.144**	0.247***	0.235***
	LL	−0.074	**−0.286*****	**−0.400*****	**−0.406*****



(a) Morphological changes (b) Lesion changes

Fig. 2. Axial (top row) and mid-sagittal (bottom row) slices of volumes representing the spectrum of MSFC scores from 1.5 to −4.5. A decrease in MSFC shows the classic pattern of MS pathology

4 Conclusions

We have introduced a new method for modeling the variability in brain morphology and lesion distribution of a large set of MRIs of MS patients. Our method is a statistical model of brain images composed of three DBNs: one for morphology, one for lesion distribution, and one that jointly models both. We have demonstrated that such a model, which requires no built-in priors on image similarity, can automatically discover patterns of variability that can be parameterized in a low-dimensional space and are clinically relevant. In addition, our model can generate sample images from model parameters for visualization. For future work, we plan to incorporate clinical parameters into the learning stage using discriminative fine-tuning, which we expect would result in further improvements in the clinical score correlations. In addition, we would like to extend our approach to model longitudinal patterns of variability with the goal of predicting future clinical status from current images.

Acknowledgements This work was supported by NSERC and the Milan and Maureen Ilich Foundation.

References

1. Ceccarelli, A., Jackson, J., Tauhid, S., Arora, A., Gorky, J., Dell'Oglio, E., Bakshi, A., Chitnis, T., Khoury, S., Weiner, H., et al.: The impact of lesion in-painting and registration methods on voxel-based morphometry in detecting regional cerebral gray matter atrophy in multiple sclerosis. AJNR American Journal of Neuroradiology 33(8), 1579–1585 (2012)
2. Wolz, R., Aljabar, P., Hajnal, J.V., Rueckert, D.: Manifold learning for biomarker discovery in MR imaging. In: Wang, F., Yan, P., Suzuki, K., Shen, D. (eds.) MLMI 2010. LNCS, vol. 6357, pp. 116–123. Springer, Heidelberg (2010)
3. Aljabar, P., Wolz, R., Srinivasan, L., Counsell, S.J., Rutherford, M.A., Edwards, A.D., Hajnal, J.V., Rueckert, D.: A combined manifold learning analysis of shape and appearance to characterize neonatal brain development. IEEE Transactions on Medical Imaging 30(12), 2072–2086 (2011)
4. Wolz, R., Aljabar, P., Hajnal, J.V., Lötjönen, J., Rueckert, D.: Nonlinear dimensionality reduction combining MR imaging with non-imaging information. Medical Image Analysis 16(4), 819–830 (2012)
5. Cayton, L.: Algorithms for manifold learning. Univ. of California at San Diego Tech. Rep., 1–17 (2005)
6. Gerber, S., Tasdizen, T., Fletcher, P.T., Joshi, S., Whitaker, R.: Manifold modeling for brain population analysis. Medical Image Analysis 14(5), 643–653 (2010)
7. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deep belief nets. Neural Computation 18(7), 1527–1554 (2006)
8. Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, A.Y.: Multimodal deep learning. In: 28th International Conference on Machine Learning, pp. 689–696 (2011)
9. Fonov, V., Evans, A.C., Botteron, K., Almli, C.R., McKinstry, R.C., Collins, D.L.: Unbiased average age-appropriate atlases for pediatric studies. NeuroImage 54(1), 313–327 (2011)
10. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17(2), 825–841 (2002)
11. Andersson, J.L.R., Jenkinson, M., Smith, S.: Non-linear registration, aka spatial normalisation. Technical Report TR07JA2, FMRIB Centre, Oxford, United Kingdom (2007)
12. Hinton, G.E.: A practical guide to training restricted Boltzmann machines. Technical Report UTML TR 2010-003, Department of Computer Science, University of Toronto (2010)
13. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y.: Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In: Proceedings of the 26th Annual International Conference on Machine Learning, pp. 609–616 (2009)
14. Fischer, J.S., Rudick, R.A., Cutter, G.R., Reingold, S.C.: The multiple sclerosis functional composite measure (MSFC): an integrated approach to MS clinical outcome assessment. Multiple Sclerosis 5(4), 244–250 (1999)
15. Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P., Federico, A., De Stefano, N.: Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. NeuroImage 17(1), 479–489 (2002)