

Combining Spatial and Non-spatial Dictionary Learning for Automated Labeling of Intra-ventricular Hemorrhage in Neonatal Brain MRI

Mengyuan Liu¹(✉), Steven P. Miller², and Vann Chau² and Colin Studholme^{1,3}

¹ Department of Bioengineering, University of Washington, Seattle, WA, USA
liumyzoe@uw.edu

² The Hospital for Sick Children, Toronto, ON, Canada

³ Department of Pediatrics, University of Washington, Seattle, WA, USA

Abstract. A specific challenge to accurate tissue quantification in premature neonatal MRI data is posed by Intra-Ventricular Hemorrhage (IVH), where severe cases can be accompanied by extreme and complex Ventriculomegaly (VM). IVH is apparent on MRI as bright signal pooling within the ventricular space in locations related to the original bleed and how the blood pools and clots due to gravity. High variability in the location and extent of IVH and in the shape and size of the ventricles due to ventriculomegaly (VM), combined with a lack of large sets of training images covering all possible configurations, mean it is not feasible to approach the problem using whole brain dictionary learning. Here, we propose a novel sparse dictionary approach that utilizes a spatial dictionary for normal tissues structures, and a non-spatial component to delineate IVH and VM structure. We examine the behavior of this approach using a dataset of premature neonatal MRI scans with severe IVH and VM, finding improvements in the segmentation accuracy compared to the conventional segmentation. This approach provides the first automatic whole-brain segmentation framework for severe IVH and VM in premature neonatal brain MRIs.

1 Introduction

In neonates born prematurely, focal pathology is one of the abnormalities that can be visualized by medical imaging techniques. Magnetic Resonance Imaging (MRI) has been an essential tool to diagnose and monitor the condition, which can be greatly improved by an accurate volumetric and morphological analysis of the neonatal brain anatomy [4], especially in cases of high-risk preterm newborns. Intra-ventricular hemorrhage (IVH) is one type of focal lesion and can often be accompanied by an enlargement of cerebral ventricles (VENT), termed Ventriculomegaly (VM)[11]. In these cases, IVH leads to errors in quantifying

Supported in part by NIH with Grants R01 NS055064, R01 NS061957 and CIHR grant.

WM and VENT volumes and shapes that are critical in basic measures of growth. Knowledge of the location and size of IVH could also reveal how IVH specifically influences development. The challenge of the automatic delineation task is two-fold. Firstly, the presence of IVH and VM makes the accurate non-rigid mapping of a normal atlas or even an IVH subject to a new IVH scan challenging because of the changes in topology required to map between the anatomy with and without regions of IVH. Secondly, it is difficult to approach the problem by building an exhaustive dictionary that collects all possible shape, size and location of IVH and the enlarged VENT.

Due to the differences in tissue contrasts in premature neonatal brain imaging, specialized atlases and methods to use and validate them have been developed [5]. Cheng et al. [2] proposed a stochastic process based approach for white matter injury detection in premature neonates. Qiu et al. [7] developed a multi-phase geodesic level-sets method that specifically targeted post-hemorrhagic ventricle dilation. However, neither method labeled the normal tissue structures in the image. Wang et al. [12] developed a patch-driven level set approach for normal term-birth neonatal T1-weighted MRIs. Liu et al. [6] proposed to integrate a local patch-based search into a spatio-temporal atlas-based method to more accurately delineate the detailed structures in pre-natal scans of varying ages. In related work, Roy et al. [8] presented a subject-specific sparse non-local dictionary learning approach for adult brain lesion segmentation. To the best of our knowledge, there has been little previous work in developing automated whole-brain segmentation methods for premature neonatal MRI scans with IVH and severe VM.

In this paper we propose to utilize a specially designed dictionary, which consists of a spatial and a non-spatial component to account for both healthy and abnormal structures. The spatial dictionary encodes normal variation in anatomy, while we use a non-spatial dictionary to capture the shape and occurrence of IVH voxels with respect to their commonly neighboring tissues. An Elastic Net algorithm is used to ensure the sparsity of the dictionary learning in both the dictionaries. The two dictionaries are collectively used to estimate a probability of normal and abnormal tissues for each voxel, which is then used to initiate an Expectation-Maximization based tissue labeling of the image data [1, 6, 10].

2 Methods

2.1 Preliminaries

The problem being addressed is to assign an initial tissue probability map to a new unseen scan. Let I be the new image under investigation, T_h be sets of lesion-free labelled MR template images $I^t (t = 1, \dots, T_h)$ with labels $L^t (t = 1, \dots, T_h)$, and T_l be sets of labelled images with IVH and VM. At voxel location x of testing image I , its intensity patch of its $p \times p \times p$ neighboring voxels is represented as a column vector Y_x , and its corresponding dictionary is denoted as D_x with size d . The sparse dictionary search task is to determine the sparse

coefficients β multiplied by which the dictionary can represent the image patch Y_x under investigation. We estimate β by solving a minimization of the non-negative Elastic-Net problem:

$$\min_{\beta, \beta \geq 0} \frac{1}{2} \|Y_x - D_x \beta\|_2^2 + \lambda_1 \|\beta\|_1 + \frac{\lambda_2}{2} \|\beta\|_2^2 \quad (1)$$

The L_1 -norm regularization ensures sparsity of β , and the L_2 -norm regularization encourages similar dictionary patches to have similar coefficients. Conventionally in brain tissue segmentation, a spatial dictionary $\overline{D_x^{sp}}$ is constructed [12] to capture locally specific information. In this work, we consider parts of the anatomy for which we do not have enough training data to capture the full range of possible locations of pathology. We propose to use an additional non-spatial dictionary $\overline{D^n}$ which is combined with the spatial dictionary, such that $\overline{D_x} = \{\overline{D_x^{sp}}, \overline{D^n}\}$. This non-spatial component is used to augment the assignment of tissue labels where abnormalities are known to occur. In the following section, we focus on the construction of our proposed combined dictionary which includes spatial samples to match the normal tissue structures such as gray matter (GM) and white matter (WM), and non-spatially encoded samples of abnormal structures, i.e. IVH and VM.

2.2 Dictionary Construction

Non-spatial Dictionary. The aim of this dictionary is to learn the appearance of focal pathologies and their occurrence with surrounding normal tissues, but to encode them without spatial constraints of where they may occur. This then can be used in regions where we assume the pathology can occur. In the problem considered here, the non-spatial component of the dictionary is constructed from ventricular regions with IVH and severe VM from lesion templates, i.e. $I^t, L^t, t = 1, \dots, T_l$. For each voxel z within this region, we extract its $p \times p \times p$ intensity patch Y_z in the form of a column vector with unit L_2 norm, arranged to form a dictionary matrix $\widetilde{D^n}$. Due to the volume of severely enlarged ventricles, the number of columns (denoted as C) of this matrix can be large ($C \sim 10^4$) with many similar columns. To reduce computation time, we remove the duplicate dictionary samples while keeping the unique ones, by thresholding the similarity measurement between samples. We define the correlation between the i -th column and j -th column of $\widetilde{D^n}$ as $corr(\widetilde{D^n}(i), \widetilde{D^n}(j))$. Then we consider the j -th column of $\widetilde{D^n}$ as a duplicate of the i -th column and remove it if

$$\max_{c, c \in [1:C]} |corr(\widetilde{D^n}(j), \widetilde{D^n}(c)) - corr(\widetilde{D^n}(i), \widetilde{D^n}(c))| < a \quad (2)$$

where a is a chosen threshold. The use of correlation mimics the patch matching criteria in the LARS sparse dictionary search process [3]. After removing the duplicate dictionary samples, we obtain a succinct non-spatial dictionary $\overline{D^n}$, which is independent of voxel location x .

Spatial Dictionary. Using a conventional approach, the spatial component of the dictionary is constructed using lesion-free templates with similar gestational ages, i.e. $I^t, L^t, t = 1, \dots, T_h$. For voxel x , we build its spatial dictionary as follows. Let \mathcal{N}_x^t denote the $N \times N \times N$ neighborhood of voxel x in t -th ($t = 1, \dots, T_h$) template image. For each voxel $z \in \mathcal{N}_x^t$, we extract its intensity patch from I^t , normalize it to have a unit L_2 norm and then rewrite it into a p^3 -sized column vector Y_z . By arranging $N^3 \times T_h$ column vectors, we obtain the spatial dictionary matrix \overline{D}_x^{SP} for each voxel x .

Combined Dictionary. For each voxel the correspondent dictionary \overline{D}_x is the combination of the spatial and non-spatial component: $\overline{D}_x = \{\overline{D}_x^{SP}, \overline{D}^n\}$. To further simplify computation, we conduct a pre-screening of the mean intensity of the dictionary patches. We exclude the dictionary patch at j -th column of \overline{D}_x if

$$\frac{|avg(\overline{D}_x(j)) - avg(Y_x)|}{avg(Y_x)} \geq b \quad (3)$$

where $avg()$ computes the mean intensity of the patch intensities before unit L_2 -normalization, b is a chosen threshold. Another benefit of the mean pre-screening is to remove the confusion caused by the dictionary sample patches with similar intensity pattern but very different absolute intensity level. For example, an uniform patch inside VENT should not be matched to the uniform patches inside WM with their similar pattern but different absolute intensity. After this, we have the final dictionary D_x for voxel x .

2.3 Implementation Details

Pre-processing. To construct the dictionary, we first linearly align all training images I and globally standardize the intensity scaling factor. For the non-spatial component, we extract voxels label as IVH or VENT (Fig. 1 (A-i)), and smoothly dilate the region to include the outer boundaries (Fig. 1 (A-ii)). An example non-spatial dictionary \overline{D}^n obtained after the duplicate removal process is shown in Fig. 1 (A-iii). A conventional atlas-based automated segmentation is used to provide outer cerebral boundary.

Sparse Dictionary Search Using LARS. For the sparse dictionary search, we use the combined dictionary for regions inside the cerebral boundary where the pathology can occur, and spatial-only dictionary for the other regions to save computation. The Elastic-Net problem (Eq. 1) is a convex optimization problem and, in our implementation, β is solved by the LARS algorithm with non-negative constraints [3]. Each element of β represents the similarity between Y_x and the corresponding dictionary sample. For LARS, the similarity is based on correlation for matching the pattern in the two patches. Under the assumption that similar patches share the same tissue label, we can compute the estimate

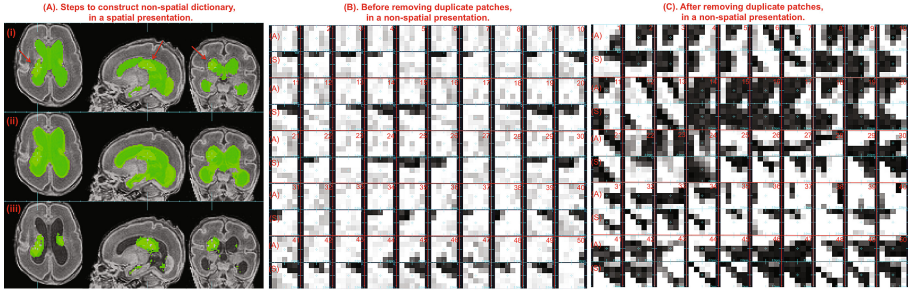


Fig. 1. Non-spatial dictionary construction. (A): Example showing the construction of non-spatial dictionary in the form of a mask (green) overlaying the subject MRI. (i) IVH and VENT mask extracted from manual labeling; (ii) dilated mask that includes duplicate voxels with similar intensity profile; (iii) remaining voxels after removing duplicates. Red arrow: IVH. (B)(C): 50 randomly selected sample patches (shown in axial and sagittal view) in the non-spatial dictionary before (B) and after (C) removing the duplicate patches. It is clearly shown that, before removal (B), more patches share the same intensity profile and will hence contribute same information to the non-spatial dictionary while unnecessarily increasing the computation time. After removal (C), we obtain more structural diversity given the same number of dictionary samples.

the tissue probability $P(k|x)$ of the voxel x belonging to tissue class k , from β as follows:

$$P(k|x) = \frac{\sum_{i=1}^d \beta_i L_i}{\sum_{i=1}^d \beta_i} \quad (4)$$

Post-dictionary EM Segmentation. The patch-based dictionary-learned tissue probability estimate is used to initiate an EM-based tissue labelling framework. The EM algorithm clusters the voxels with similar intensities into same tissue classes given the prior tissue estimates. The final automated tissue labeling is obtained. Our full segmentation driven has in total 8 labels. In the following section, we focus only on the 5 cerebral tissue structures that contain lesions: GM, WM, VENT, deep gray matter (DGM) and IVH.

3 Experimental Results

3.1 Dataset and Validation

Our data consists of a total of 12 T1-weighted MR scans of premature neonatal brains with manual tracing into GM, WM, VENT, DGM, cerebellum (CBL), brain stem (BS), sulcal CSF (sCSF) and IVH, 4 of which have IVH and severe VM. To test the approach we used 2 age matched normal scans to construct a spatial dictionary and 3 out of 4 of the IVH scans to construct the non-spatial dictionary leaving one to be automatically segmented. This was repeated for each

Table 1. Gestational ages of testing datasets and the corresponding dictionary data.

Test Case	Test Scan	Spatial Dict. (2 scans)		Non-spatial Dict. (3 scans)		
1	28.7 GW	28.6 GW	29.4 GW	29.4 GW	32.7 GW	33.6 GW
2	29.4 GW	28.6 GW	29.4 GW	28.7 GW	32.7 GW	33.6 GW
3	32.7 GW	32.6 GW	33.1 GW	28.7 GW	29.4 GW	33.6 GW
4	33.6 GW	33.1GW	33.7 GW	28.7 GW	29.4 GW	32.7 GW

of the 4 IVH cases and Dice Similarity Coefficients (DSC) calculated against the corresponding manual label. The experimental data is summarized in Table 1.

3.2 Parameter Selection

Optimal parameter values were determined by leave-one-out cross-validation on all 4 IVH scans. Values for L_1 -regularization coefficient $\lambda_1 = 0.1, 0.2, 0.3, 0.4, 0.5$, patch size $p = 3, 5$ and neighborhood size $N = 3, 5, 7$ were compared using DSC (Fig. 2), finding optimal values: $\lambda_1 = 0.2, p = 5, N = 7$. We also tested on a smaller scale and chose L_2 -regularization coefficient $\lambda_2 = 0.01$ as in [12], dictionary thresholds $a = 0.04$ and $b = 0.2$. The impact of duplicate dictionary sample removal using $a = 0.04$ is shown in Fig. 1(B)(C).

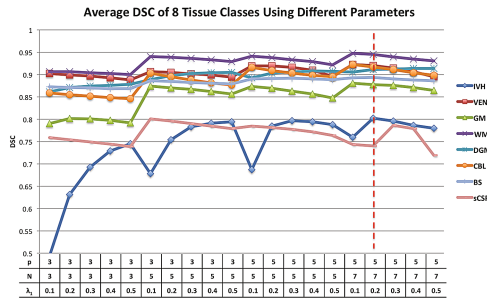


Fig. 2. Average DSC of 8 tissue classes (IVH, GM, WM, VENT, DGM, BS, CBL and sCSF) with respect to the different combinations of parameters λ_1, p and N .

3.3 Results

To show the contribution from the spatial and non-spatial components of the dictionary, we compare the number of positive coefficients in the spatial and non-spatial part of β for each voxel in Fig. 3 (A), confirming that for normal tissue the primary contribution is from the spatial dictionary, while locations with abnormal ventricles or IVH are determined by the non-spatial dictionary. To show the effect of the EM algorithm, we compared the before- and after-EM DSC average: IVH: 0.6931 to 0.8129; VENT: 0.8385 to 0.9321; GM: 0.7853 to

0.8780; WM: 0.9021 to 0.9474; DGM: 0.8852 to 0.9116. We can clearly see that dictionary labelling provides an accurate initial tissue label estimate and then EM adapts it further by modeling of subtle residual bias field to improve the final labels.

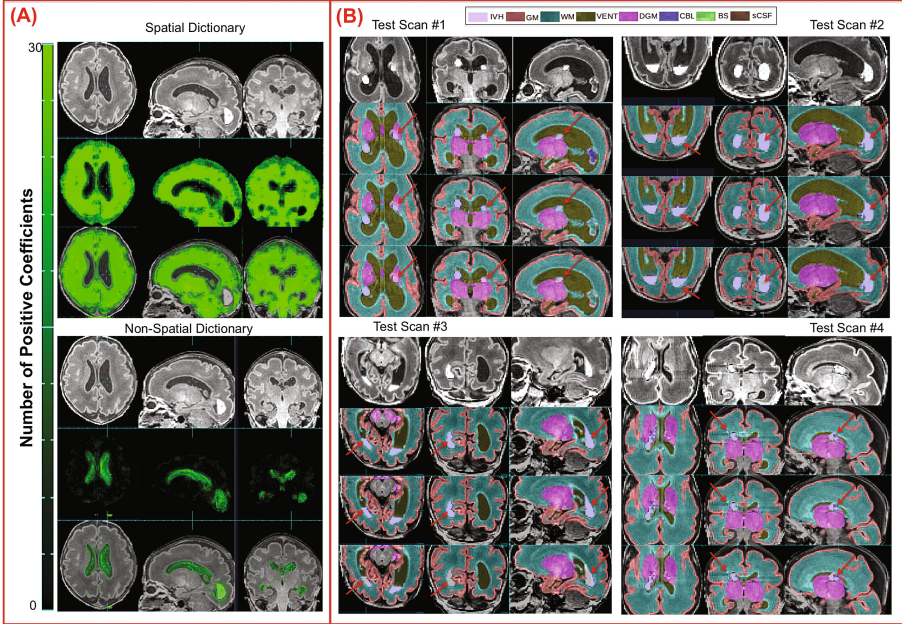


Fig. 3. (A) Number of positive similarity coefficients from spatial (top) and non-spatial (bottom) dictionaries for each voxel. (B) Comparison of manual (2nd row), automatic using spatial and non-spatial dictionary (3rd row) and automatic using only spatial dictionary (4th row) tissue segmentation of all 4 testing scans. Red arrow: the IVH region that is correctly labeled using the proposed combined dictionary while mislabeled using spatial-only dictionary.

Table 2. Comparison of individual and average DSC of 5 main tissue classes obtained by using the proposed combined dictionary (left section) and spatial-only dictionary (right section).

Scan	Spatial + Non-spatial Dictionary					Spatial-only Dictionary				
	IVH	VENT	GM	WM	DGM	IVH	VENT	GM	WM	DGM
1	0.6522	0.9524	0.8849	0.9520	0.8951	0.4833	0.9608	0.8872	0.9599	0.9405
2	0.8813	0.9386	0.8885	0.9377	0.9313	0.9234	0.9509	0.9047	0.9554	0.9429
3	0.8746	0.9144	0.8660	0.9434	0.8982	0.8373	0.9209	0.8672	0.9435	0.9064
4	0.8436	0.9230	0.8724	0.9564	0.9219	0.8329	0.9247	0.8815	0.9589	0.9262
Avg	0.8129	0.9321	0.8780	0.9474	0.9116	0.7692	0.9393	0.8852	0.9544	0.9290

Figure 3 (B) summarizes the key results with red arrows indicating where the combined dictionary improved performance. In particular scan #1 in Fig. 3 (B) illustrates a case where the IVH location was not present in the training data leading to a failure in the spatial-only dictionary approach, but a correct labelling when also using the non-spatial dictionary. Table 2 summarizes the average DSC scores confirming the overall improvement in IVH segmentation across the IVH cases.

4 Conclusion

This paper describes a novel hybrid technique to address segmentation of highly variable focal abnormalities that is motivated by the study of abnormally developing premature neonatal brain anatomy. The proposed method seeks to label brain anatomy with a tissue probability using a collective sparse search of a combined spatial and a non-spatial dictionary to provide a more accurate estimate of the tissue labels, for both focal lesions and surrounding normal tissues. The spatial component represents the normal anatomical variations and the non-spatial component encodes the variable appearance of IVH and VM. Experimental analysis of the results of EM segmentation driven by this prior, compared against manually delineated premature neonatal brain MRIs indicated improved performance. Future work entails adaption of a discriminative dictionary learning approach [9] for dictionary construction, to further distinguish IVH from intraparenchymal blood, and to carry out an extensive validation in other age ranges when data becomes available.

References

1. Ashburner, J., Friston, K.J.: Unified segmentation. *Neuroimage* **26**(3), 839–851 (2005)
2. Cheng, I., Miller, S.P., Duerden, E.G., Sun, K., Chau, V., Adams, E., Poskitt, K.J., Branson, H.M., Basu, A.: Stochastic process for white matter injury detection in preterm neonates. *NeuroImage Clin.* **7**, 622–630 (2015)
3. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al.: Least angle regression. *Ann. Stat.* **32**(2), 407–499 (2004)
4. Hüppi, P.S., Warfield, S., Kikinis, R., Barnes, P.D., Zientara, G.P., Jolesz, F.A., Tsuji, M.K., Volpe, J.J.: Quantitative magnetic resonance imaging of brain development in premature and mature newborns. *Ann. Neurol.* **43**(2), 224–235 (1998)
5. İşgum, I., Benders, M.J., Avants, B., Cardoso, M.J., Counsell, S.J., Gomez, E.F., Gui, L., Hüppi, P.S., Kersbergen, K.J., Makropoulos, A., et al.: Evaluation of automatic neonatal brain segmentation algorithms: the neobrain12 challenge. *Med. Image Anal.* **20**(1), 135–151 (2015)
6. Liu, M., Kitsch, A., Miller, S., Chau, V., Poskitt, K., Rousseau, F., Shaw, D., Studholme, C.: Patch-based augmentation of expectation-maximization for brain MRI tissue segmentation at arbitrary age after premature birth. *NeuroImage* **127**, 387–408 (2016)

7. Qiu, W., Yuan, J., Rajchl, M., Kishimoto, J., Chen, Y., de Ribaupierre, S., Chiu, B., Fenster, A.: 3D MR ventricle segmentation in pre-term infants with post-hemorrhagic ventricle dilatation (phvd) using multi-phase geodesic level-sets. *NeuroImage* **118**, 13–25 (2015)
8. Roy, S., Carass, A., Prince, J.L., Pham, D.L.: Subject specific sparse dictionary learning for atlas based brain MRI segmentation. In: Wu, G., Zhang, D., Zhou, L. (eds.) *MLMI 2014*. LNCS, vol. 8679, pp. 248–255. Springer, Cham (2014). doi:[10.1007/978-3-319-10581-9_31](https://doi.org/10.1007/978-3-319-10581-9_31)
9. Tong, T., Wolz, R., Wang, Z., Gao, Q., Misawa, K., Fujiwara, M., Mori, K., Hajnal, J.V., Rueckert, D.: Discriminative dictionary learning for abdominal multi-organ segmentation. *Med. Image Anal.* **23**(1), 92–104 (2015)
10. Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imaging* **18**(10), 897–908 (1999)
11. Volpe, J.J.: Intraventricular hemorrhage in the premature infant current concepts. part i. *Ann. Neurol.* **25**(1), 3–11 (1989)
12. Wang, L., Shi, F., Li, G., Gao, Y., Lin, W., Gilmore, J.H., Shen, D.: Segmentation of neonatal brain MR images using patch-driven level sets. *NeuroImage* **84**, 141–158 (2014)