# A challenging issue: Detection of white matter hyperintensities in neonatal brain MRI

Baptiste Morel[1,2], Yongchao Xu[2,3], Alessio Virzi[2], Thierry Géraud[3], Catherine Adamsbaum[2,4], Isabelle Bloch[2]

1. Faculty of Medicine, François Rabelais University, Pediatric Radiology, CHRU Tours, France

2. LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, Paris, France

3. EPITA Research and Development Laboratory (LRDE), Le Kremlin-Bicêtre, France

4. Faculty of Medicine, Paris Sud University, Pediatric Radiology Department, Bicêtre Hospital APHP, Paris, France

***Abstract*— The progress of magnetic resonance imaging (MRI) allows for a precise exploration of the brain of premature infants at term equivalent age. The so-called DEHSI (diffuse excessive high signal intensity) of the white matter of premature brains remains a challenging issue in terms of definition, and thus of interpretation. We propose a semi-automatic detection and quantification method of white matter hyperintensities in MRI relying on morphological operators and max-tree representations, which constitutes a powerful tool to help radiologists to improve their interpretation. Results show better reproducibility and robustness than interactive segmentation.**

## I. Introduction

Prematurity concerns 15 millions of children every year. Newborn brain MR images provide qualitative and quantitative data about the maturation of the brain, such as myelination [4]. The controversy about the prognosis of the diffuse excessive high signal intensity, initially described in [3] then in [14], remains actual and debated, e.g. [5], [7], [10]. However, the detection (and the reproducibility of the comparisons) of high signal intensity in the white matter by radiologists constitutes a difficult task, prone to subjectivity and intra and inter-observer variability [8].

This issue led us to propose a robust image processing method to provide a more reproducible and objective analysis of the signal intensity of the white matter in neonates. Actually, a semi-automatic tool to detect and quantify the hyperintensities in the white matter would constitute an important help to determine their definition and thereby their prognosis. However neonatal MRI images show significant differences with respect to older children or adult brain images, such as inverted contrast and small dimensions, and existing segmentation methods cannot be applied. Therefore, we have developed a new pipeline for the segmentation of different neonatal brain tissues (grey and white matter, cerebrospinal fluid, grey nuclei) on 1.5T T2-weighted newborn brain images, 1.5T still being the most used magnetic field in clinical routine. Concerning white matter hyperintensities, a lot of methods have been developed for adult brain images [1]. For instance some algorithms are based on an optimal FLAIR intensity threshold to separate the hyperintensities from normal brain tissues, based on the analysis of image histograms [6]. Samaille et al. have developed an unsupervised method relying on a non-linear diffusion filter and including the local image contrast as an important parameter [13]. But, to the best of our knowledge, no semi-automatic nor automatic method for detecting white matter intensities is available for newborn brain images. In this paper we focus on the detection and quantification of hyperintensities in such images. The proposed method relies on mathematical morphology based on max-tree representations, using image contrast as in [13], but it is specifically designed for neonatal brain images.

## II. Overview of the segmentation method

Brain tissues are extracted step by step from the original images. We first apply an anisotropic diffusion filter [11] to decrease the noise. Then the suppression of the scalp is realized by opening and image thresholding. The intracranial cavity (ICC) is then obtained, in which all next processing steps are restricted. The following steps are based on a max-tree representation [12] (i.e. a hierarchical representation of the connected components of all upper level sets of the image, the parenthood relation being defined by inclusion). Indeed, in T2 images of neonatal brains, cephalo-spinal fluid (CSF), ventricles and hyperintensities are brighter than their surroundings, and thus appear in the max-tree representation as nodes or branches that can be clearly identified.

To segment the CSF (in the pericerebral spaces and within the ventricles), markers are first defined by thresholding the image (default value is set to 0.85 for an image normalized between 0 and 1). The connected components in the max-tree that include the markers and that satisfy a contrast criterion are then selected. Pericerebral CSF and lateral ventricles (LV) are then separated based on the location of markers.

For the extraction of the basal ganglia and thalami (BGT), an area closing is first applied with a large area value (default value is set to 2/3 of the ICC size) and the original image is subtracted from the closing result. This provides an image in which the BGT is again a region brighter than its surroundings, and that can therefore be detected in the max-tree representation. Additionally to markers inside the BGT, the detection is also constrained by an enclosing rectangle that defines the region of interest. A regularization step based on maximal Cheeger set is then applied [2], to smooth the contours.

When subtracting all detected structures from the ICC, the remaining tissues are composed of the cortical grey matter

(GM) and the white matter (WM), including the hyperintensities. The grey levels of these tissues are separated using a simple histogram analysis, leading to an optimal threshold using Otsu's method [9]. To avoid a potential important overlap between the grey levels distributions of grey matter and white matter on the whole brain parenchyma, we propose to perform the histogram analysis in small rectangular blocks partitioning the ICC.

Interactive correction tools are offered to the user at each step for potential local improvements if needed.

## III. Segmentation and quantification of hyperintensities

Within the white matter, hyperintensities are then segmented. Again, the white matter hyperintensities are brighter than their surroundings, hence correspond to regions in the max-tree representation. A context-based energy is used to define a contrast criterion. Assuming that the intensity function $f$ in a uniform region $R$ can be approximated by its average intensity value $\overline{f}(R)$, the quantity

$$V(f,R) \;=\; \sum_{p\in R} (f(p) - \overline{f}(R))^2$$

can be interpreted as a segmentation error. Let $\partial R$ be the boundary of region $R$, we define the interior context region $\mathcal{R}^{\varepsilon}_{in}(\partial R)$ and exterior context region $\mathcal{R}^{\varepsilon}_{out}(\partial R)$ as the sets of points at a distance less than $\varepsilon$ from $\partial R$, respectively inside and outside the region $R$. We use the measurement defined as:

$$E(f,\partial R) \;=\; \frac{V(f,\, \mathcal{R}^{\varepsilon}_{in}(\partial R)) \,+\, V(f,\, \mathcal{R}^{\varepsilon}_{out}(\partial R))}{V(f,\, \mathcal{R}^{\varepsilon}_{in}(\partial R) \cup \mathcal{R}^{\varepsilon}_{out}(\partial R))}. \quad (1)$$

This measurement is between 0 and 1, and is low when the segmentation error is much lower for two classes than for one class, meaning that the curve is an object contour. See [15], [16] for more details about the context-based energy and its computation. The regions in the max-tree representation that have a small context-based energy value defined in Equation (1) correspond to the tissues to be extracted. Usually too many nodes may be selected, and markers can then be helpful, as used in the previous steps (Section II). Unfortunately for white matter hyperintensities, inside markers are difficult to obtain. Therefore we propose a selecting and discarding process to extract the objects, relying on the fact that the hyperintense regions are disjoint connected components. More precisely, we first spot the "most likely" region $R^*_1$ among all the regions in the tree, based on the energy value, and discard all the ancestors and descendants of $R^*_1$. Then we retrieve a second "most likely" region $R^*_2$ among the remaining regions in the tree, and discard again its descendants and ancestors. This selecting and discarding process is repeated until all the regions are either spotted or discarded. Finally, a set of regions $\{R^*_i, i = 1,\ldots,n\}$ are spotted, where the number of spotted objects $n$ is decided by the algorithm. Only the regions having a relatively small energy value (e.g. smaller than 0.5) and a high average grey level intensity (e.g. larger than 0.8 for an image normalized between 0 and 1) are considered as candidate hyperintensity regions.

For each candidate hyperintensity region, we compute the relative difference of grey levels compared to its surrounding white matter region. A hyperintensity region should have a high relative difference. We use this criterion to further filter out some candidate hyperintensity regions. Note that some manual corrections can be applied easily to refine the extraction of white matter hyperintensities.

## IV. Results and conclusion

We have developed a graphical interface integrating each step of the pipeline, which makes the step by step application of all operations easy. We have also integrated some manual corrections that enable the user to improve the segmentation results for each tissue without much effort. The manual corrections are summarized as follows: remove a whole existing region by selecting an inside point; remove part of an existing region by drawing a closed contour; add a region by drawing a closed contour; select a region by drawing a closed contour; select an existing region by choosing an inside point. These manual corrections (usually only very few are needed) are easy to perform, and help to improve the segmentation results.

To validate the robustness of the detection of white matter intensities, we have selected ten axial T2 weighted images of premature newborns (born between 28 and 29 weeks of gestation) whose clinical status, transfontanellar ultrasound, and electroencephalogram were all normal. The images were acquired at term-equivalent age (between 39 and 40 weeks of gestation) on a 1.5 MRI device with an 8-element head coil. The acquisition parameters were as follows: TR = 3750 ms; TE = 110 ms; Turbo-factor = 16. The MRI slices were acquired without interpolation, with pixel size from $0.36 \times 0.36$ to $0.7 \times 0.7 mm^2$. An example is shown in Figure 1. Pixel size varies from $0.36 \times 0.36\, mm^2$ to $0.7 \times 0.7\, mm^2$ depending on the cases. Slice thickness varies from 2 to $4\, mm$. Due to the high anisotropy of the data, the whole processing was applied on 2D slices.

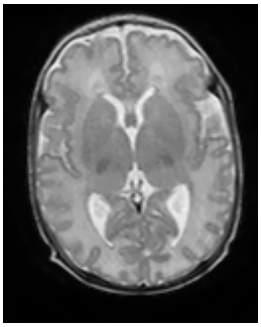

Fig. 1. One axial slice of a T2 MRI volume of a premature newborn brain.

One senior observer has manually segmented the different brain tissues. Two senior observers have manually detected and segmented potential white matter hyperintensities to

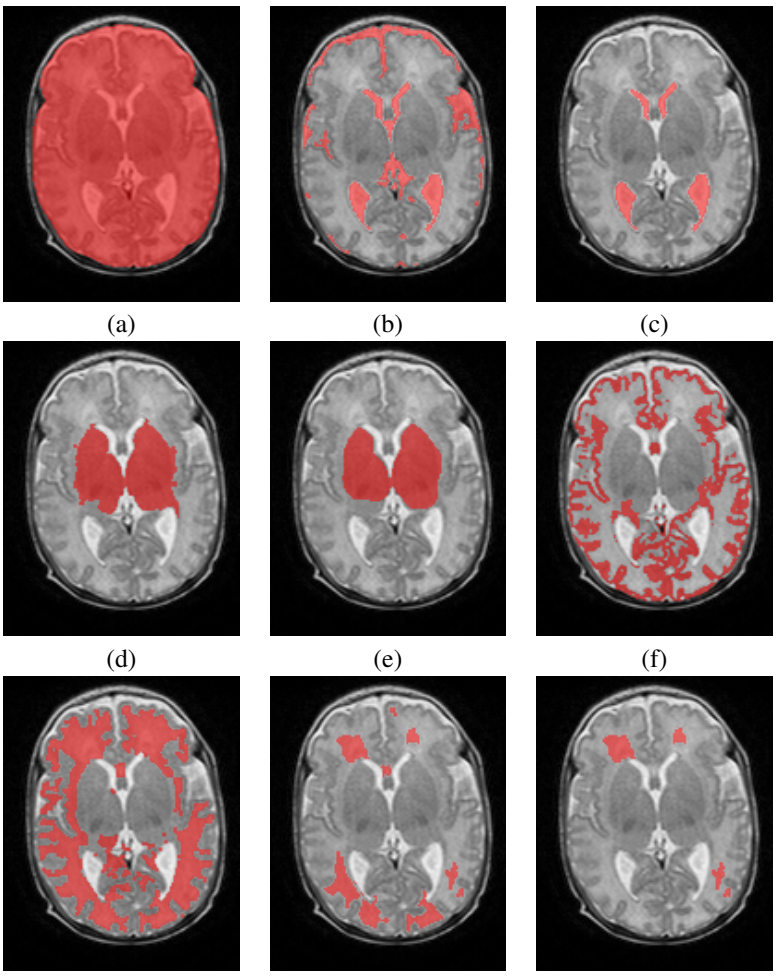


Fig. 2. Segmentation results on the example in Figure 1. (a) ICC. (b) CSF and ventricles. (c) LV. (d) BGT before regularization. (e) BGT after regularization). (f) GM. (g) WM. (h) Candidates for hyperintensities. (i) Selected candidates for hyperintensities. Each structure is superimposed in red on the original slice.

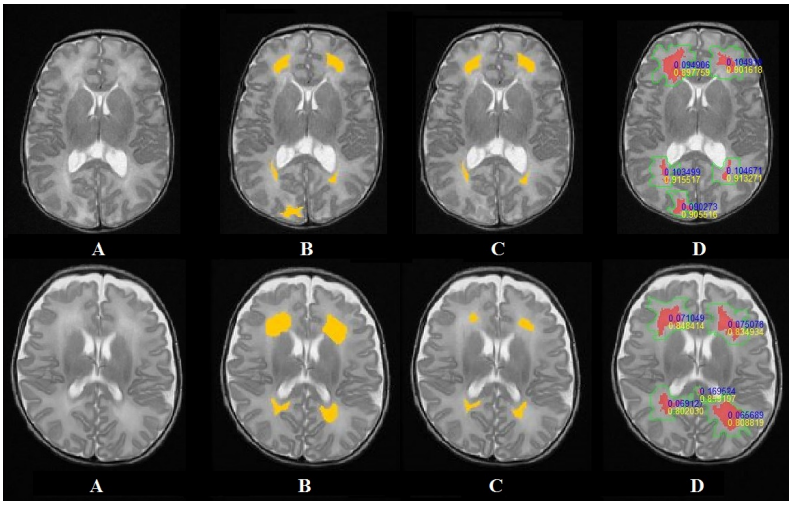


Fig. 3. Hyperintensity detection for two patients. (A) original image. (B, C) Manual detection of potential white matter intensity by two different senior observers. (D) Semi-automatic detection with the proposed method. Values in yellow and blue indicate the normalized average intensities in the hyperintensities (red regions) and the relative difference of grey levels compared to their surroundings in the white matter (delimited by the green contours).

assess the inter-individual variability. Because of the few potential manual corrections applied during the semi-automated segmentation, the variability of the use of the software was assessed by two successive applications of the method with one month delay. Dice similarity indices were calculated to evaluate the quality and reproducibility of the segmentation of the algorithm.

Results on one case are illustrated in Figure 2 on the example in Figure 1. Visually results are good. It should be noted that precise boundaries are difficult to determine on such images because of the partial volume effect induced in particular by the slice thickness.

Table I provides the results, averaged over all cases, for the main brain structures (Section II), comparing the results of the proposed method (Auto.) and of the manual segmentation by the senior observer (Obs.), and two usages of the automatic method (Auto.1 and 2), with potentially different manual interactions. It shows that the results of the proposed method are good, with generally high Dice similarity index. The fact that this index is quite sensitive with respect to shape (smaller indices are generally observed on small or thin structures) may explain the smaller values obtained for the grey matter, which is anyway known to be very difficult to observe on such images, in particular due to the strong partial volume effect on thin structures. Moreover the manual segmentation was sometimes incomplete. The stability of the algorithm according to the potential manual corrections is excellent.

TABLE I

AVERAGE DICE SIMILARITY INDEX FOR THE DIFFERENT BRAIN STRUCTURES.

| Structure | Obs. vs. Auto. | Auto.1 vs. Auto.2 |
|---|---|---|
| CSF | 0.76 | 0.96 |
| LV | 0.85 | 0.99 |
| BGT | 0.95 | 0.99 |
| GM | 0.59 | 0.97 |
| WM | 0.75 | 0.94 |

Table II provides results for the hyperintensities. As shown in [8], deciding visually whether hyperintensities are present or not is a highly subjective matter, and the precise delineation of the corresponding regions is even more difficult. This is confirmed by the lower values of dice index. However, the algorithm remains very reproducible, which is a main improvement over manual detection, and the variability with respect to observers is of the same order as the variability between observers. More qualitatively, all hyperintensities detected by the senior observer are also detected by the automatic method, and no additional one was found. Only the limits and spatial extent can be detected differently. These results are illustrated visually for two patients in Figure 3.

These preliminary results demonstrate the potential of the proposed semi-automatic method, with very limited user interactions, for segmenting brain structures and indicating potential hyperintensities in standard neonatal 1.5T MRI imaging data. The high reproducibility of the algorithm is an important improvement over existing visual assessment methods. It will thereby constitute a powerful tool with useful clinical applications.

TABLE II

AVERAGE DICE SIMILARITY INDEX FOR THE DETECTIONS OF POTENTIAL HYPERINTENSITIES.

| Obs.1 vs. Obs.2 | Obs.1 vs. Auto. | Obs.2 vs. Auto. | Auto.1 vs. Auto.2 |
|---|---|---|---|
| 0.49 | 0.42 | 0.49 | 0.98 |

## REFERENCES


[1] M. E. Caligiuri, P. Perrotta, A. Augimeri, F. Rocca, A. Quattrone, and A. Cherubini. Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review. *Neuroinformatics*, 13(3):261–276, 2015.

[2] G. Carlier, M. Comte, and G. Peyré. Approximation of maximal Cheeger sets by projection. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43(01):139–150, 2009.

[3] S.J. Counsell, J.M. Allsop, M.C. Harrison, D.J. Larkman, N.L. Kennea, O. Kapellou, F.M. Cowan, J.V. Hajnal, A.D. Edwards, and M.A. Rutherford. Diffusion-weighted imaging of the brain in preterm infants with focal and diffuse white matter abnormality. *Pediatrics*, 112(1):1–7, 2003.

[4] S.J. Counsell, E.F. Maalouf, A.M. Fletcher, P. Duggan, M. Battin, H.J. Lewis, A.H. Herlihy, A.D. Edwards, G.M. Bydder, and M.A. Rutherford. MR imaging assessment of myelination in the very preterm brain. *Am. J. Neuroradiology*, 23(5):872–881, 2002.

[5] F.T. de Bruïne, A.A. van den Berg-Huysmans, L.M. Leijser, M. Rijken, S.J. Steggerda, J. van der Grond, and G. van Wezel-Meijler. Clinical implications of MR imaging findings in the white matter in very preterm infants: a 2-year follow-up study. *Radiology*, 261(3):899–906, 2011.

[6] C.R. Jack, P.C. OBrien, D.W. Rettman, M.M. Shiung, Y. Xu, R. Muthupillai, A. Manduca, R. Avula, and B. J. Erickson. FLAIR histogram segmentation for measurement of leukoaraiosis volume. *J. Magnetic Resonance Imaging*, 14(6):668–676, 2001.

[7] T.Y. Jeon, J.H. Kim, S.-Y. Yoo, H. Eo, J.-Y. Kwon, J. Lee, M. Lee, Y.S. Chang, and W.S. Park. Neurodevelopmental outcomes in preterm infants: comparison of infants with and without diffuse excessive high signal intensity on MR images at near–term-equivalent age. *Radiology*, 263(2):518–526, 2012.

[8] B. Morel, G. Antoni, J.P. Teglas, I. Bloch, and C. Adamsbaum. Neonatal brain MRI: how reliable is the radiologists eye? *Neuroradiology*, pages 1–5, 2015.

[9] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[10] N.A. Parikh, R. E. Lasky, K.A. Kennedy, G. McDavid, and J.E. Tyson. Perinatal factors and regional brain volume abnormalities at term in a cohort of extremely low birth weight infants. *Plos One*, 2013.

[11] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[12] P. Salembier, A. Oliveras, and L. Garrido. Antiextensive connected operators for image and sequence processing. *IEEE Trans. Image Processing*, 7(4):555–570, 1998.

[13] T. Samaille, L. Fillon, R. Cuingnet, E. Jouvent, H. Chabriat, D. Dormont, O. Colliot, and M. Chupin. Contrast-based fully automatic segmentation of white matter hyperintensities: method and validation. *PloS one*, 7(11):e48953, 2012.

[14] L.J. Woodward, P.J. Anderson, N.C. Austin, K. Howard, and T.E. Inder. Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. *New England Journal of Medicine*, 355(7):685–694, 2006.

[15] Y. Xu, E. Carlinet, T. Géraud, and L. Najman. Efficient computation of attributes and saliency maps on tree-based image representations. In *Int. Symp. Mathematical Morphology*, volume 9082 of *LNCS*, pages 693–704, 2015.

[16] Y. Xu, T. Géraud, and L. Najman. Context-based energy estimator: Application to object segmentation on the tree of shapes. In *IEEE Int. Conf. Image Processing*, pages 1577–1580, 2012.