On the Longitudinal Impact of Exposure Bias in Recommender Systems

Automated morphological measurements of brain structures and identification of optimal surgical intervention for Chiari I malformation

Luca Mesin, Forough Mokabberi, and Christian Francesco Carlino

***Abstract*—The herniation of cerebellum through the foramen magnum may block the normal flow of cerebrospinal fluid determining a severe disorder called Chiari I Malformation (CM-I). Different surgical options are available to help patients, but there is no standard to select the optimal treatment. This paper proposes a fully automated method to select the optimal intervention. It is based on morphological parameters of the brain, posterior fossa and cerebellum, estimated by processing sagittal magnetic resonance images (MRI). The processing algorithm is based on a non-rigid registration by a balanced multi-image generalization of demons method. Moreover, a post-processing based on active contour was used to improve the estimation of cerebellar hernia. This method allowed to delineate the boundaries of the regions of interest with a percentage of agreement with the delineation of an expert of about 85%. Different features characterizing the estimated regions were then extracted and used to develop a classifier to identify the optimal surgical treatment. Classification accuracy on a database of 50 patients was about 92%, with a predictive value of 88% (tested with a leave-one-out approach).**

***Index Terms*—Magnetic resonance imaging, demons, non-rigid registration, active contour, Chiari malformation.**

I. Introduction

CHIARI I Malformation (CM-I) arises from the development of cerebellar tonsil, i.e., the descent of the lower part of the brain through the foramen magnum [1]. The herniated tissue, extending down through the skull and into the spinal canal, could block the normal flow of cerebrospinal fluid (CSF). As a consequence, fluid may buildup in the brain (hydrocephalus) or in the spinal cord (syringomyelia), depending on the site of blockage.

CM-I is probably determined by an underdevelopment of the para-axial mesoderm of the posterior fossa (i.e., deficient bone growth during fetal development with consequent poor expansion of the posterior fossa and content overcrowding) [2], but other forms have also been reported: neuroectodermal (i.e., defect of neural tube during fetal development, associated to more severe Chiari malformations, often incompatible with life) and acquired forms (i.e., cerebellar herniation due to brain masses that cause intracranial hypertension) [3]. The herniation of the cerebellar tonsils can show varying degrees of symptomatology which may include suboccipital headaches, occipital pain exacerbated by coughing or by the Valsalva maneuver, ataxia, dysphagia, sensory deficits, scoliosis and muscle weakness [4].

There are few data on the progression of symptoms in asymptomatic patients. Moreover, it is not so easy to establish whether a cephalalgic symptomatology is linked to CM-I or to an alternative etiology, such as migraine or tension headache.

A tonsillar ectopy of 5 mm was adopted as a "cut-off" for the diagnosis of CM-I [4] [5] [6] [7]. However, there is a growing recognition that the extent of herniation is not a good indicator of the severity of the symptoms or of the clinical evolution [8]. Indeed, in recent years, with the increased availability of magnetic resonance images (MRI), the number of incidental or asymptomatic CM-I has increased [9]. Moreover, some patients have cerebellar tonsils, but no symptoms (Chiari 0 malformation, even if some neurosurgeons have expressed their objections regarding the existence of this type of CM [10]). This finding has also questioned the role of tonsillar herniation in the severity of the pathology [11].

There are also variations in the assessment of CM-I in dependence of the age of the patient. In children, tonsillar ectopia is very common (about 78% of brain scans with MRI), but many of them (about 37%-57%) are asymptomatic at the time of diagnosis [12]. On the other hand, only 14%-30% of adult patients are asymptomatic to radiological diagnosis [13]. This is consistent with the observation that CM-I symptoms take time to develop, often becoming clinically evident at about the age of 30-40 years. Indeed, the cerebellar tonsils typically rise with age [14]. Siringomyelia is also more frequent in adults (59%-76%) than in children [15].

For these reasons, there are many controversies regarding the choice of the correct surgical treatment [16] [17], which should expand the capacity of the posterior fossa and establish a normal CSF flow through the foramen magnum and foramen of Magendie [18]. The following options are available.

- Posterior fossa decompression (PFD), which is the primary treatment for children and adults with symptoms referable to the malformation. The traditional technique for decompression involves suboccipital craniectomy with C1 laminectomy, dural opening, and duraplasty with or without coagulation of cerebellar tonsils.
- PFD without dural opening, which seems also to be sufficient to arrest the progress of the disease and to improve the preoperative symptomatology in a high percentage

L. Mesin is with the Mathematical Biology and Physiology group, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy e-mail: luca.mesin@polito.it

F. Mokabberi is with the Mathematical Biology and Physiology group, Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy.

C.F. Carlino is with the Ospedale Infantile Regina Margherita, Department of Pediatric Neurosurgery, Turin, Italy.

Manuscript received October 30, 2019; revised ...

of cases [19] [20] [21], with lower complications rate, shorter operative time and hospital stay, lower cost for patients and hospitals compared to the surgical procedure with dural opening [22].
- As a third possibility, no surgical treatment could be applied.

The correct choice of the treatment is of great importance. Indeed, the symptomatic recurrence rate after the first surgery is high (about 35% [23]). Many studies seem to confirm that only patients clearly symptomatic and/or with syringomyelia can benefit from surgery. Now, only some neurosurgeons (about the 46%) recommend the surgical intervention for a patient with occipital headache as the sole symptom and in the absence of syringomyelia [24].

Also MRI CSF flow data are not useful for the management of CM-I, because both symptomatic and asymptomatic patients show abnormal flow characteristics at the cranio-cervical junction [25]. Thus, many neurosurgeons and experts are now looking for a new way to define CM-I, to measure severity and to predict what kind of patients will develop symptoms or undergo surgical treatment.

Few studies have been proposed to find new parameters (alternative to the dimension of the cerebellar hernia) correlated to CM-I [26]. Morphometric skull analysis is one of the most active areas of research. Particularly, MRI has begun to be used to measure quantitative morphometric properties [27] [28] [29] [30]. However, measurements have been taken manually and only the discrimination of controls and patients was addressed.

The aim of this paper is to support the selection of the proper surgical intervention for CM-I by a fully automated approach with low computational cost (thus, with the aim of providing a fast feedback to the surgeon). Morphometric indexes are estimated by an automated processing of MRI and are then used to identify the best surgery approach by a machine learning algorithm.

The morphology is assessed by estimating in sagittal MRIs the regions corresponding to brain, posterior fossa and cerebellum, and extracting some indexes characterizing their shapes and relative dimensions. Different approaches have been proposed in the literature to label different portions of the brain [31], e.g., Statistical Parametric Mapping [32] and registration over an atlas [33]. A non-rigid registration was here considered. The simple demons approach was selected (instead of more sophisticated methods, e.g., based on statistical models [34] or deep learning [35] [36]), as it has been appreciated for its good performances and low computational cost [37]. A post-processing was also needed, as, for topological reasons, cerebellar herniation could not be properly estimated if the atlas and test images were not compatible: the problem could be possibly solved by using a multi-atlas approach [38], but, in order to keep a low computational cost, the problem was managed by an adjustment of region boundaries by an active contour method [39].

Morphometric analysis was applied to a database of MRIs from controls and CM-I pediatric patients. The patients were divided into 3 classes, based on the symptoms and the type of optimal surgical treatment, selected a-posteriori on the basis of the applied treatment and outcome. A machine learning algorithm was applied to identify the correct surgery procedure, given the morphometrics extracted from the MRI of each subject. Many artificial intelligence approaches have been developed and applied in biomedical applications [40], but, due to our small dataset and for the sake of simplicity, standard methods were here considered (i.e., Naïve Bayes and binary tree models[1]). Different classifiers were fit to the dataset, selecting the solution with best performances.

This study is an extension of a previous work [41], in which preliminary results were shown (the dataset was here extended and the processing method was improved and tested).

II. Methods

A. Preliminary estimation of brain regions by demons

Different regions in the brain were identified by the atlas approach, i.e., by registering the test image on a reference one (see Figure 1), in which brain, posterior fossa and cerebellum were delineated by an expert (an image of our dataset from a patient with a quite long cerebellar hernia was chosen as atlas). All images were first registered by an affine transformation in order to have the axes between nasion and inion with equal orientation and length. Original images were converted into gray-scale and adjusted to span all 256 levels. Then, the background was identified and put at the lowest level.

Demons method was then used for non-rigid image registration [42]. Some variations of the classical approach were also considered, by applying a method recently proposed to improve the performance of demons, but still keeping a low computational cost [41] [43]. Specifically, a multi-image generalization of demons was introduced. The multiple images were produced by applying to each MRI different filters or texture operators. The same deformation was assumed to apply to each image pair. Moreover, the algorithm was implemented balanced between the two images, instead of considering a "fixed" and a "moving" image, as done by the classical approach, which deforms only the latter. In the following, a short introduction to demons is given first, then the multi-image generalization and the balanced approach are described.

1) Classical demons: Demons method is a non-parametric, non-rigid image registration based on a diffusion process [44]. To match two images, object boundaries in one image are considered as semipermeable membranes through which the other image diffuses thanks to the action of effectors (called demons, situated within those membranes). Diffusion is driven by forces depending on polarity, i.e., by intensity variations across these object boundaries. In practice, MRI intensity gradient determines the normal to an object boundary and the deformable model is pushed in such a normal direction in order that the two images are matched [37].

Consider two images to be registered, defined as functions of each pixel p: the fixed $F(p)$ and the moving $M(p)$ images. The registration requires to identify the deformation $s : p \rightarrow s(p)$ such that the deformed image $M \circ s(p)$ fits the reference image $F(p)$, in terms of a similarity functional. The mean squared error is considered as functional to be minimized.

[1]Additional tests using k Nearest Neighbor and a set of SVMs provided poorer results.

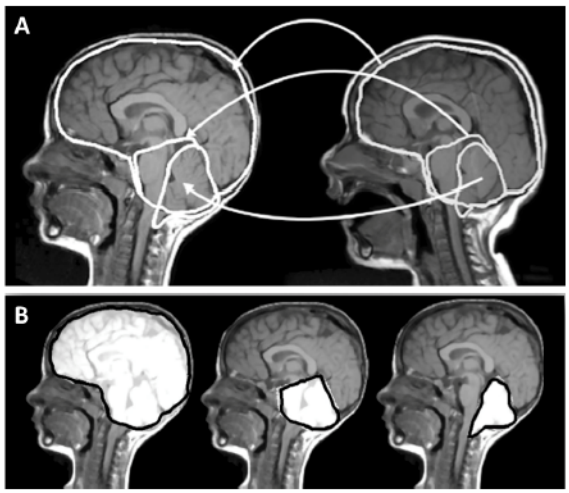

Fig. 1. A) Delineation of brain regions by the atlas approach: the test image (left) is compared to the atlas (right), in which the regions were drawn by an expert. B) Comparison between the regions delineated manually and identified automatically on the test image: from left to right, brain, posterior fossa and cerebellum are considered; light regions drawn by the expert, black curves computed automatically.

Moreover, a regularization term is added, to stabilize the solution. Furthermore, to improve computational efficiency [42], a hidden variable c (of correspondences) is introduced, requiring to be an approximation of the transformation s accommodating possible errors. Finally, the following functional was minimized with respect to s and c

$$\frac{\|F(p) - M \circ c(p)\|^2}{\sigma_i^2} + \frac{\|c(p) - s(p)\|^2}{\sigma_x^2} + \frac{\|\nabla s(p)\|^2}{\sigma_T^2} \quad (1)$$

where σ_i, σ_x and σ_T are parameters accounting for image noise, spatial uncertainty in the correspondences and regularization, respectively.

This problem was solved iteratively: in each step, the sum of the first two terms was minimized with respect to c with s fixed; then the sum of the last two terms was minimized with respect to s with c fixed. The latter step simply requires to compute a Gaussian smoothing of c (a standard deviation of 1.5 pixels was considered in this paper). The first step was approached by solving iteratively local linearization problems till convergence to a solution. Writing $c = s \circ (\mathbb{1} + u)$, where $\mathbb{1}$ is the identify map and u is a small deformation, the following update for the deformation was obtained [43]

$$u = \frac{(F - M \circ s) J_p^T}{\|J_p\|^2 + \frac{\sigma_i^2}{\sigma_x^2}} \quad (2)$$

J_p being the Jacobian operator in the point p. The following choice of regularization was considered [43]

$$\frac{\sigma_i^2}{\sigma_x^2} = |F - M \circ s|^2 \quad (3)$$

2) Multi-image generalization: Demons algorithm was generalized to estimate a single deformation allowing to match optimally different images from the atlas and the test MRI. Those images can be obtained either by processing the original one with filters or extracting texture features. In this way, information from a neighbor of each pixel is summarized in each location. The deformation optimally matching the images obtained from the atlas and the test MRI by the same processing procedures is assumed to be the same as the one matching the original images. This is only approximately true, as the neighbor considered for the processing is affected by the deformation, introducing a perturbation with effects which are not trivial to predict. However, tests in simulations showed that the multi-image approach, by averaging different information included in each image, allows to estimate better a known deformation [43].

Following the results of [43], the multi-image approach was applied to the following 5 images: the original one, the contrast-limited adaptive histogram equalized [45] (with diameter of the neighbors equal to 32 pixels), the image after application of a local median filter (with neighborhood of diameter 5 pixels), the local image entropy (with circular neighborhood of radius 15 pixels) and the phase symmetry of the image [46].

The method consists in the following procedure. Assume that N images $\{M_i\}$ obtained from the test subject are to be mapped to an atlas represented by N images $\{F_i\}$. The optimization problem to be solved is obtained by summing more mean squared errors, one for each image pair. The local displacement field u can be obtained analytically [43]

$$u = \frac{\sum_{i=1}^{N} (F_i - M_i \circ s) J_i^T}{\sum_{i=1}^{N} \|J_i\|^2 + \Sigma_i^2} \quad (4)$$

which is the multi-image generalization of (2), where $\Sigma_i^2 = |F_i - M_i \circ s|^2$. The estimated displacement field was then smoothed and applied to the moving map to update it. The transformation was imposed to be a diffeomorphism [42] (i.e., a smooth and invertible map). Specifically, the update was applied through the exponential map, i.e., instead of computing $c = s \circ (\mathbb{1} + u)$ as above, the deformation was updated as $c = s \circ exp(u)$.

As mentioned above, tests in simulations showed that this multi-image algorithm has better performance than the classical demons in the estimation of a simulated deformation [43]. The computational cost is slightly increased (the computational time was about the 70% larger), as more images have to be processed.

3) Balanced algorithm: The standard algorithm deforms only one of the two images in a pair (the moving test image is deformed over the fixed atlas, which is not deformed). However, in this way, the test image is perturbed by the interpolation method used to apply the deformation. As an alternative, a balanced algorithm was implemented that, at each iteration, estimates and applies deformations in alternation both to the atlas and to the test images. Notice that the theoretical solution of the problem is not unique, as, given the two estimated deformations to be applied to the (sets of) images, an additional arbitrary deformation could be composed to both of them. However, a unique numerical solution was achieved. Finally, the deformation that transforms

the test image into the atlas is given by the composition of the direct deformation starting from the original test image and the inverse of the deformation of the atlas (which can be computed, as diffeomorphisms are considered).

Tests in simulations showed that this method allows to further improve the estimation of the deformation [43]. Moreover, as indicated in [43], this method performs slightly better than a symmetric approach [47] [48], computing the update as an arithmetic average of forward and backward transformations (i.e., mapping the moving into the fixed image and vice-versa). It requires only a small additional computational cost with respect to the non-balanced version: the computational time for the balanced multi-image algorithm was about the 115% larger than that of the standard demons.

B. Correction of region boundaries by active contour

The registration algorithm allows to get only a preliminary identification of the main brain regions of interest and in particular of the cerebellum. For example, topological reasons could prevent the possibility of estimating correctly the cerebellar tonsil. In fact, consider a test subject without the cerebellar hernia and an atlas showing it. As the estimated transformation is diffeomorphic, if the hernia is present in the atlas, it will be identified also in the test image. In fact, under our hypothesis, the cerebellum of the atlas has a portion which is out of the posterior fossa; this portion cannot be mapped by a continuous transformation inside the fossa (as it would be required to match perfectly the cerebella of the atlas and test images).

A correction of the boundaries of the regions estimated by the registration algorithm was then implemented. Around the regions of interest, there are some portions of the MRI with low intensity. For example, around the cerebellum, there are the cerebellar tentorium and the portion at the bottom which have low intensity (Figure 2A). However, those regions are not all around the cerebellum: for example, the boundary separating it from the brainstem is not clearly visible. Thus, the boundaries of the cerebellum had to be updated by a displacement toward these not connected low intensity regions around it and the other portions of the boundary should be obtained by interpolation (and the same holds for the posterior fossa and the brain). The boundary was also assumed to be smooth.

An active contour (or snake) was used to update the boundaries [39]. It can be interpreted as an elastic string with a certain rigidity reflected by internal constraints (enforcing it to have a smooth shape) and pushed by external forces (which displace it toward the regions of interest). The active contour has a dynamics dictated by the following energy functional to be minimized

$$E(\mathbf{v}) = \int_0^1 \left[E_{int}(\mathbf{v}(s)) + E_{ext}(\mathbf{v}(s))\right] ds \qquad (5)$$

where $\mathbf{v}(s) = [v_x(s)\,,\ v_y(s)]$ is the parametric representation of the active contour (i.e., a 1D curve, depending on the curvilinear abscissa s) and E_{int} and E_{ext} are the internal and external energies, respectively. The active contour evolves searching for a local minimum of the energy functional (5). Its specific final configuration will depend on the initial condition (that in our case is the boundary identified by the registration algorithm) and the energies, detailed below.

- The internal energy was defined as [39]

$$E_{int}(\mathbf{v}(s)) = \left(\alpha \left|\mathbf{v}_s(s)\right|^2 + \beta \left|\mathbf{v}_{ss}(s)\right|^2\right)/2 \qquad (6)$$

  where $\mathbf{v}_s$ and $\mathbf{v}_{ss}$ are the first and second derivatives of $\mathbf{v}(s)$, respectively; α=0.2 controls the tension and β=1 is related to the rigidity of the active contour (a fine tuning based on a few preliminary tests was applied to select the values of parameters).
- The external forces push the contour toward the region of interest (thus, toward low intensity points in the image). Specifically, the test image was binarized by two-class k-means clustering, in order to identify the low intensity regions[2]. A further step was required to process the cerebellum. Indeed, the cerebellum has similar gray-scale intensity as the brainstem and medulla; thus, the border of such a region was identified and converted into black, so that the active contour could be attracted by it[3]. The distance transform from the dark pixels was then computed, by assigning to each point the distance from the closest dark pixel of the test image. The distance transform provided the potential function which was used to define the external forces as [39]

$$E_{ext} = w_{line} E_{line} + w_{edge} E_{edge} \qquad (7)$$

  where E_{line} is the distance transform of the binarized test image mentioned above and E_{edge} is minus the square of the modulus of the gradient of the distance transform, w_{line}=10 and w_{edge}=1 (more importance was given to approaching dark regions than edges; the parameters were chosen after preliminary fine tuning on a few MRIs).

Figure 2 shows the correction of the estimation of the borders of the cerebellum by the active contour approach.

C. Tests

1) Experimental data: Sagittal T1-weighted head MRI images at 1.5 Tesla were taken from 50 patients (28 pediatric, 22 control adults). Patients received different kinds of treatments:

[2]The background was converted into white, so that it could not attract the active contour.

[3]The following procedure was adopted to estimate the boundary of the brainstem and the medulla (Figure 2D). The left border of brainstem was identified by demons and was followed in the downward direction till going about 10 cm under the posterior fossa, into the neck. Then, fixing the row of the image, the right border of medulla was estimated searching for the first local minimum of intensity in the right direction. A portion of the right border of medulla was then estimated by following upward the dark line delimiting it, up to a few cm from the posterior fossa (in order to be sure to be under a possible hernia). The upper part of the right border of the brainstem was well estimated as the left border of the cerebellum identified by demons (the three fourth of the upper part of the left border of the cerebellum was used, leaving undetermined only the last portion, possibly including the hernia). Now, the estimated right borders of the brainstem and of the medulla were interpolated with a quadratic polynomial in order to define the interface between cerebellum and medulla in the intermediate region. Then, the dark regions were updated including also this interface.

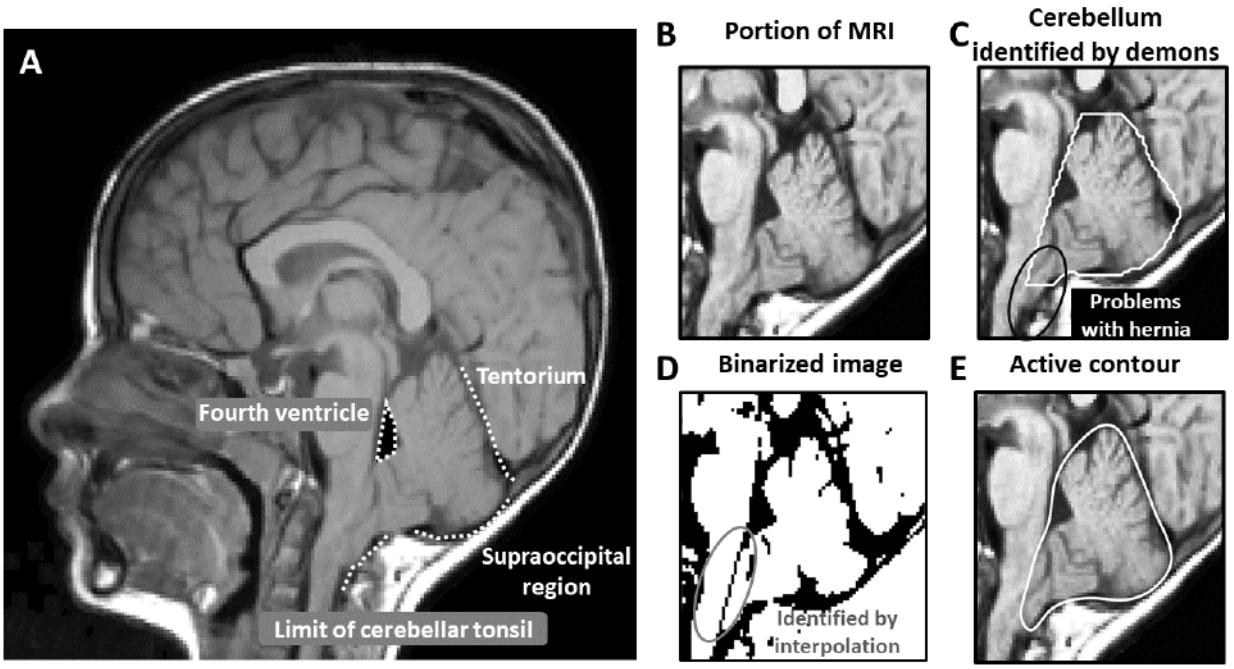

Fig. 2. Correction of the borders of the cerebellum by active contour. A) Test image with indication of some dark regions. B) Portion of the image around the cerebellum. C) Delineation obtained by the balanced multi-image demons. D) Binarized image (including the border between cerebellum and brainstem obtained by interpolation). E) Correction of the delineation of the cerebellum obtained by active contour.

no treatment, posterior fossa decompression, durotomy, tonsils coagulation and duraplasty. On the basis of their outcome (e.g., improvement of the symptoms or complete resolution; either good or bad recovery), they were split into the following classes:

1 Healthy (32 subjects; they should not be treated);
2 Mild disease (10 subjects; they may be initially treated with posterior fossa decompression without durotomy);
3 Severe condition (8 subjects; they should be surgically treated with durotomy and duraplasty, with or without tonsils coagulation).

2) Automated identification of the class: Different classifiers were tested to identify the class of the patients from morphometric measurements extracted from the automated detection of the brain regions.

Two different classification approaches were tested to fit our multi-class problem (both implemented in MATLAB R2019a[4]): the Naïve Bayes model (NBM) and the binary tree model (BTM) [49].

- The development of a NBM requires to fit the data with specific probability distributions. A smoothing density estimation with normal kernel was used.
- The BTM was implemented using the Gini's diversity index as splitting criterion. The best categorical predictor split was chosen from all possible combinations of choices.

[4]The Mathworks, Natick, Massachusetts, USA

They were applied on the following different features extracted from the identified regions including the brain, the posterior fossa and the cerebellum:

1) percentage ratio between the surface areas (in pixels) of the regions including the cerebellum and the posterior fossa;
2) percentage ratio between the areas of the cerebellum and of the brain;
3) length of cerebellar hernia (defined by the maximal distance of the boundary of the cerebellum from the foramen magnum) divided by the major axis of the cerebellum;
4) the area of the cerebellar hernia;
5) the area of the cerebellar hernia divided by the area of the cerebellum;
6) the minor divided by the major axis of the cerebellum;
7) the length of the tentorium (the border of the posterior fossa was approximated by a pentagon, from which different parameters were extracted, including the tentorium and the height of the fossa, mentioned below);
8) the length of the tentorium divided by the major axis of the posterior fossa;
9) height of the posterior fossa divided by its major axis.

The cascade of two classifiers was used, as it showed best performances in [41]. The first classifier discriminated among healthy subjects and patients, the second one identified the severity of patients. Different combinations of the above mentioned 9 features were considered as inputs to build

different classifiers (exhaustive search, considering all possible combinations of 5 features). The performances of the models were tested by cross-validation, considering 10 folds. The ones providing best generalization (i.e., minimum average mean squared regression error on the validation sets) were then selected. Due to the small dataset and the high number of tested classifiers, cross-validation is likely to provide overoptimistic performances [50]. The classification performances of the best models were then assessed with a leave-one-out approach.

D. Summary of the method

Our method is based on the following steps.

- Registration by balanced multi-image demons of a sagittal MRI of the patient, to make a preliminary estimation of the 3 regions of interest: brain, posterior fossa and cerebellum.
- Improvement of the estimation of the edges of the 3 regions by active contour, using an external force driving the region boundary toward dark regions.
- Automated measurement of morphometrics from the estimated regions.
- Application of a classifier to identify the patients to be treated. Only to those patients, a second classifier is applied to estimate if they are mild or severe (needing simple posterior fossa decompression or durotomy plus duraplasty, respectively).

III. Results

Figure 3 shows examples of MRIs processed by 4 registration methods: standard demons, multi-image algorithm, balanced multi-image and balanced multi-image corrected by active contour. The accuracy of these methods is tested comparing the manual delineations by an expert with the automated estimations in the entire dataset. The new algorithms improve the estimation of the brain regions. The active contour is statistically useful only to improve the delineation of cerebellum (due to the important correction of the cerebellar tonsil). When pooling together all estimation errors corresponding to the three regions, the median was always decreasing by considering more advanced processing methods. High statistical difference was indicated by the paired Wilkoxon signed rank test when comparing the multi-image algorithm with standard demons and the balanced multi-image with the multi-image algorithm; no statistical difference was found between including or not the active contour correction to the balanced multi-image algorithm (indeed, it was statistically useful only for correcting cerebellum).

The estimation of the cerebellar hernia is further deepened in Figure 4. Notice the great improvement in estimating the length and area of the hernia when using the active contour correction.

To further test the accuracy of our segmentation algorithm, a comparison with a convolutional neural network with UNET topology was provided in [43] considering a subset of test images. Such a deep learning approach has shown outstanding results in image processing applications [51], so that it could provide an important reference. Different architectures were tested, but the best performances were obtained considering an encoder depth of 4 (resulting in 58 layers), a cross entropy loss with learning rate of 0.1, processing images after histogram equalization. The network was trained on 40 images and tested on the remaining 10. The performances were worse than when using the proposed registration method. Specifically, as indicated in [43], the balanced multi-image approach obtained lower overlapping errors when applied to the same 10 images. Moreover, problems were found in estimating the hernia (possibly, a loss functional weighting more the region around the herniation could provide some improvement; as an alternative, the active contour discussed here could compensate for this problem, as it improved also the output of the balanced multi-image method).

Given the morphometrics extracted from the brain segmentation obtained by our method, the best classification performances were obtained using the cascade of a NBM to discriminate between healthy/asymptomatic subjects and patients and a BTM to identify the severity of patients. The first classifier includes 5 features: the ratio of the areas of cerebellum and brain, the length of the hernia over the major axis of the cerebellum, the ratio between minor and major axes of cerebellum, the length of tentorium and the length of tentorium divided by the major axis of the posterior fossa. They include information on all 3 regions investigated. Special attention is given to the dimension of cerebellar tonsils under the foramen magnum (which optimally discriminates our healthy subjects and patients, in line with literature [26]). However, other morphometric parameters reflecting the relative dimension and shape of cerebellum were also included, together with information on tentorium. The second model includes only one feature: the ratio between the height of the posterior fossa and its major axis. Thus, the shape of the posterior fossa appears to be the best feature to discriminate mild and severe patients. The confusion matrix of a leave-one-out test is shown in Table I, where the estimated classes are compared to the correct ones. The classification was correct in the 88% of cases (6 mistakes out of 50; some other common performance indexes, related to percentages of true/false positive/negatives, are provided in the caption of the table). Problems were found in discriminating the symptomatic patients (the number of available symptomatic patients is very small).

The distributions of the features for patients in different classes are shown in Figure 5. The features selected by the different classifiers are indicated. Notice that the dimension of the hernia is useful to discriminate the healthy and pathological patients, whereas the morphology of the posterior fossa is useful to discriminate different patients.

IV. Discussion

Choosing the correct surgical treatment for CM-I patients is still difficult [17]. The recurrence of symptoms after the first surgery is still quite common (about 35% of cases [23]). The literature is only focused on the identification of asymptomatic patients [26] and no objective method has been proposed to choose the best surgical option for symptomatic CM-I.

We have addressed this problem in a preliminary study, in which an innovative automated classification was proposed to

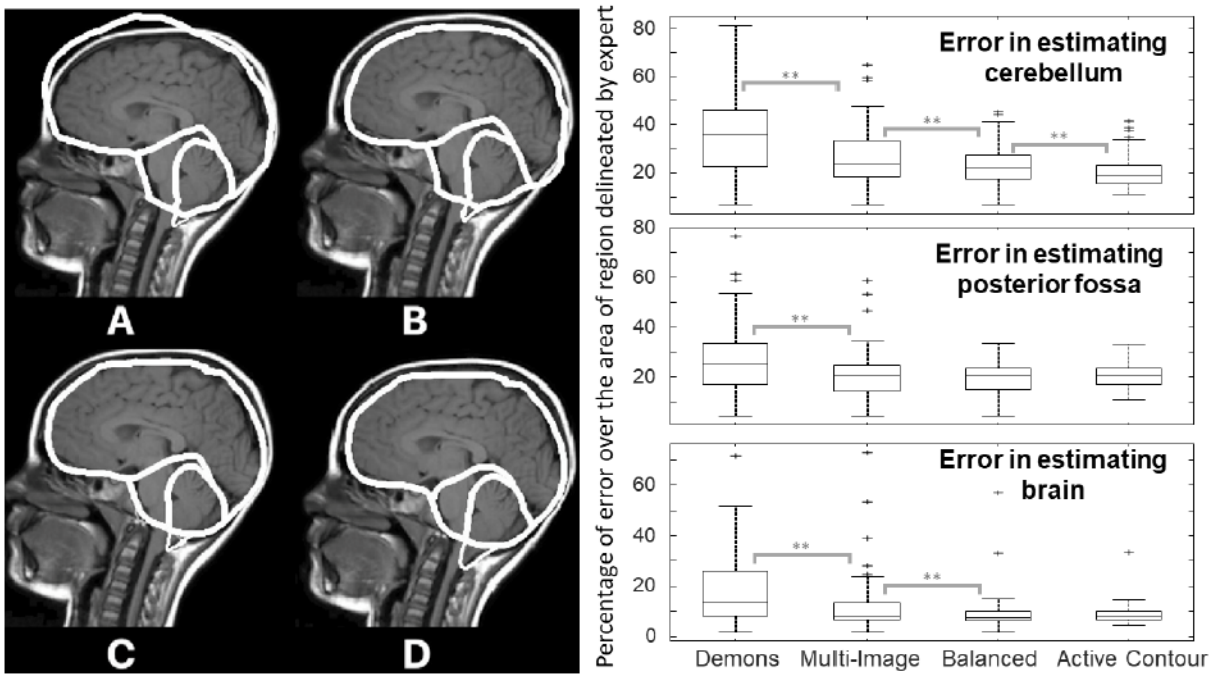

Fig. 3. Delineation obtained by different methods (left): A) standard demons, B) multi-image algorithm, C) balanced multi-image and D) balanced multi-image corrected by active contour. On the right, percentage errors (median, quartiles, range and outliers shown individually) for the entire dataset (erroneous pixels over the total number of pixels covering the regions delineated by the expert) for the same methods (** indicates highly significant differences by paired Wilkoxon signed rank test).

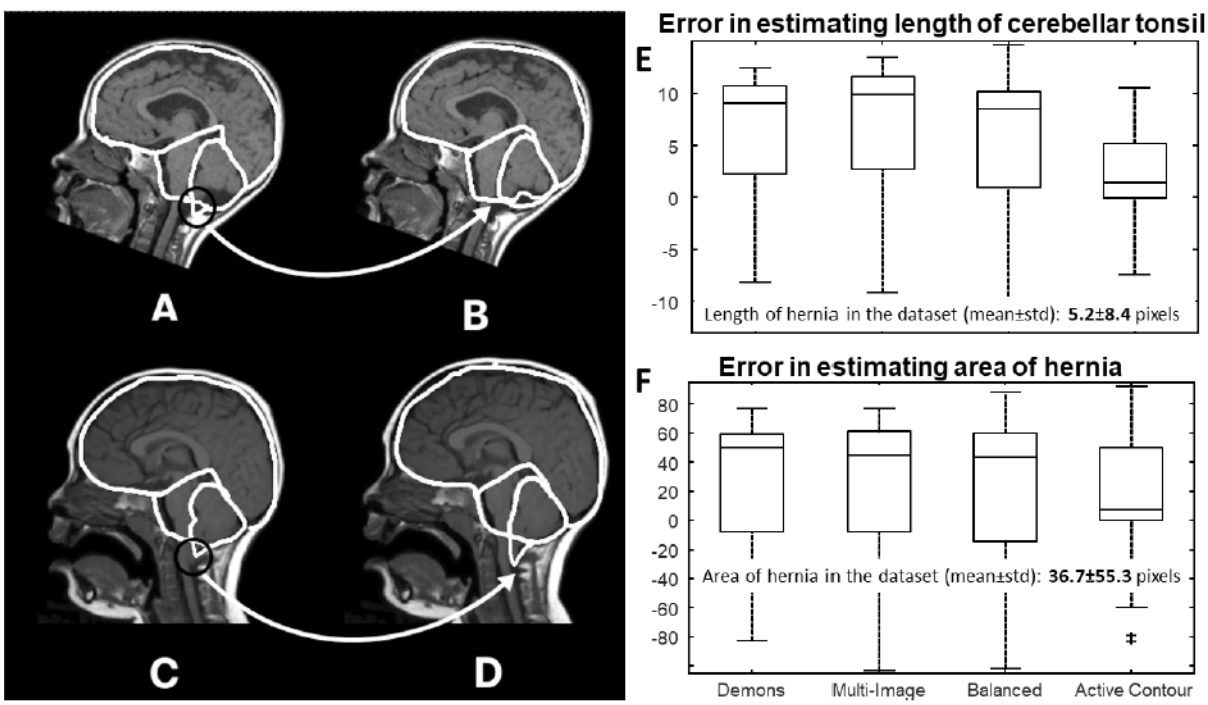

Fig. 4. Correction of cerebellar hernia using the active contour. A) Patient without the hernia delineated by the balanced multi-image method. B) Correction of A) using the active contour. C) Patient with a long hernia, underestimated by the balanced multi-image method. D) Correction of C) using the active contour. E) Distributions of errors in estimating the length of hernia in the entire database. F) Distributions of errors in estimating the area of hernia in the entire database.

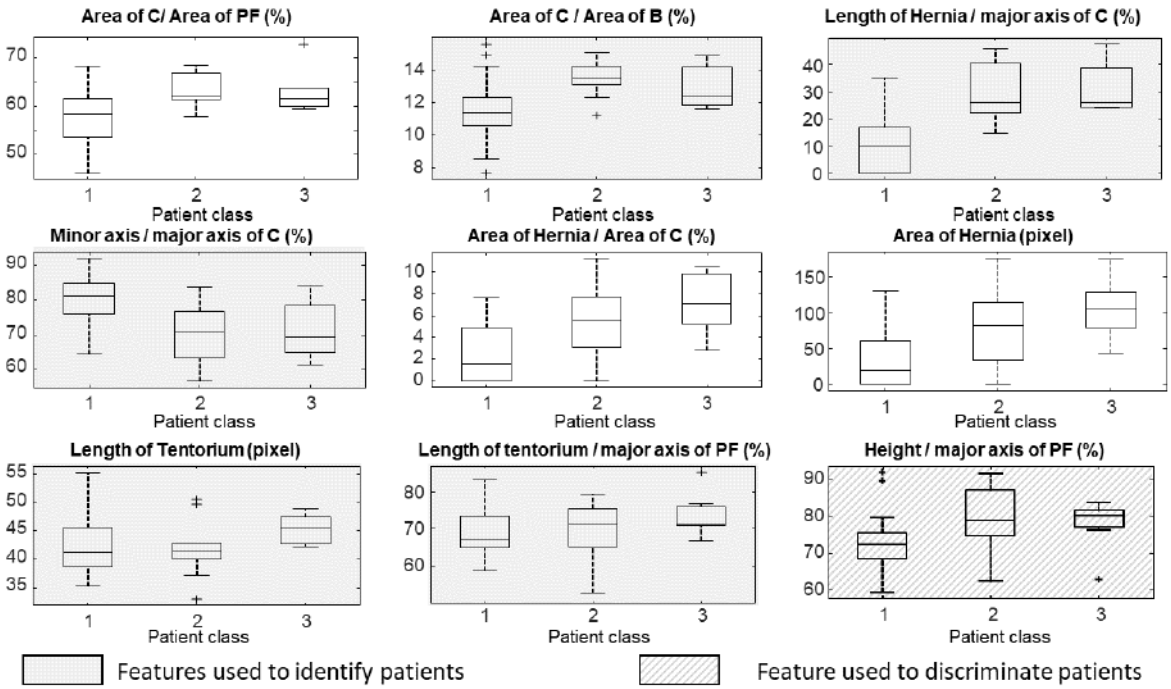

Fig. 5. Features considered to train the classifiers. The ones indicated in gray were chosen by the best NBM used to identify the symptomatic patients. The height of the posterior fossa divided by its major axis was used by the BTM that best discriminated the severity of symptomatic patients. Acronyms: B - brain; PF - posterior fossa; C - cerebellum.

TABLE I
CONFUSION MATRIX FOR THE CASCADE OF TWO CLASSIFIERS, THE FIRST SEPARATING HEALTHY SUBJECTS FROM PATIENTS, THE SECOND IDENTIFYING THE SEVERITY OF THE PATHOLOGY (LEAVE-ONE-OUT APPROACH; MEAN SENSITIVITY 88.0%; MEAN SPECIFICITY 84.8%; NEGATIVE PREDICTIVE VALUE 97.5%; MEAN ACCURACY 91.9%).

Predicted class	Target score			Predictive value
	1: Healthy	2: Mild	3: Severe	
1	32 (64%)	3 (6%)	1 (2%)	88.9%
2	0	7 (14%)	2 (4%)	77.8%
3	0	0	5 (10%)	100%
True rate	100%	70.0%	62.5%	**88.0%**

identify the severity of CM-I on the basis of morphological parameters extracted from sagittal MRIs [41]. Here, we have deepened this study. A larger dataset is considered, including 50 patients (whereas only 30 patients were discussed in [41]). For each patient, the optimal surgical intervention was determined evaluating the clinical outcomes. Data were then split into 3 classes: healthy, mild or severe patients.

Moreover, the automated method to delineate the brain, posterior fossa and cerebellum was improved (also based on a recent study focused on the registration method [43]) and its performances were assessed. Specifically, the accuracy in the automated estimation of the three regions (using the delineation by an expert as ground truth) increases when using a multi-image approach instead of the classical demons algorithm; a further improvement was obtained using a balanced algorithm, in which both images are deformed during registration. The processing method is still very fast (as it requires about twice the processing time of the classical demons algorithm), giving the possibility of providing rapid feedback to the clinician. Consider also that a diffeomorphic registration cannot fit topologically different conditions, e.g., the case in which the hernia is present in the atlas, but not in the test image. Thus, the registration should provide only a good starting point for an algorithm for boundary correction (and the use of a more sophisticated registration method, e.g., [52] [53] [54], is not required). An active contour approach was used to achieve such a correction of the delineation of the borders, especially around the cerebellar tonsil. The active contour post-processed the borders obtained by registration, largely improving the estimation of the hernia. Finally, the median overlapping error in estimating the different regions is about 20% of the pixels, for the cerebellum and posterior fossa (which are the most difficult regions to delineate). However, consider that the manual delineation is subjective and, as shown in [43], by repeating it more times, an overlapping pairwise difference of about 8% is obtained, so that optimizing further the automated segmentation could be of marginal utility. To investigate further the possible usefulness of improving the accuracy of the estimation of the brain regions in our application, the classification was also trained and tested by using the manual delineation to compute morphometrics (as if a perfect registration approach was applied, so as no mistake was made with respect to our gold standard). The identification of the correct classes of the patients did not improve under

such ideal conditions (still getting 6 misclassifications in a leave-one-out test).

In our approach, the estimation of morphometrics was completely automated, as an important difference with respect to the main literature on Chiari malformation, where usually brain regions are delineated manually [26]. Once obtained morphometric parameters, they were used to develop a classifier. The best one was based on a cascade of a NBM, which identified the symptomatic patients, and a BTM, that discriminated patients into mild or severe. Different morphological features were used by the NBM, reflecting shape and relative dimensions of the considered brain regions (including in particular the dimension of cerebellar hernia). The BTM included a single feature, reflecting the shape of the posterior fossa (which is indeed very important to characterize CM-I patients [26], and also to determine the severity of the pathology, once it is known that the patient is symptomatic). The classification accuracy slightly improved with respect to the results shown in [41]. However, some symptomatic patients have been still misclassified. Indeed, the discrimination power of the considered features was not so high (maximal Fisher discrimination ratio was about 0.3). Thus, there is the need of finding new features that better characterize the severity of the pathology, in order to be able to better identify the correct treatment. Moreover, the dimension of the dataset was increased with respect to our previous work [41] mostly by adding asymptomatic patients (for which the classification performances largely improved): more symptomatic patients of different known severity should be included in the future to further refine the ability to discriminate them.

V. Conclusion

A fully automated method is proposed to support surgeons to select the correct treatment for CM-I patients. It is based on the estimation of morphometrics characterizing different brain regions and a classification algorithm. Demons non-rigid registration is used for delineating different brain regions on sagittal MRIs. By analyzing together more images (all extracted by pre-processing the two images to be registered), the registration improves with respect to the standard demons approach, still keeping a low computational cost. Further improvement is documented in the case of a balanced algorithm. To further improve the accuracy to identify specific details, such as the cerebellar tonsil, an active contour is introduced to correct the preliminary delineation obtained by the registration. From the automated estimation of the brain, posterior fossa and cerebellum, different morphometric parameters are extracted. Some of them are used to feed a classifier that is able to correctly identify the severity of CM-I in about 90% of tested patients. Better accuracy is found in the identification of asymptomatic subjects than in discriminating the severity of symptomatic patients (possibly reflecting the scarcity of data from those patients). Thus, future investigation is needed using a larger dataset to confirm our promising results.

Acknowledgments

The authors are grateful to Dr. P. Peretta who provided access to the clinical data discussed in this work.

References

[1] P. Steinbock, "Clinical features of Chiari I malformations," *Child Nerv Syst.*, vol. 20, pp. 329–331, 2004.

[2] T.H. Milhorat, M. Nishikawa, and R.W. Kula, "Mechanisms of cerebellar tonsils herniation in patients with Chiari malformation as guide to clinical management," *Acta Neurochir.*, vol. 152, pp. 1117–27, 2010.

[3] E. Schijman, "History, anatomic forms, and pathogenesis of Chiari I malformations," *Child Nerv Syst.*, vol. 20, no. 5, pp. 323–328, 2004.

[4] T.H. Milhorat, M. Chou, and E. Trinidad, "Chiari I malformation redefined: clinical and radiographic findings for 364 symptomatic patients," *Neurosurgery*, vol. 44(5), pp. 1005–17, 1999.

[5] A.O. Aboulezz, K. Sartor, C.A. Geyer, and M.H. Gado, "Position of cerebellar tonsils in the normal population and in patients with Chiari malformation: a quantitative approach with MR imaging," *J Comput Assist Tomogr.*, vol. 9, pp. 1033–1036, 1985.

[6] B.J. Lawrence, A. Urbizu, P.A. Allen, F. Loth, R.S. Tubbs, A.C. Bunck, J.-R. Kroger, B.G. Rocque, C. Madura, J.A. Chen, M.G. Luciano, R.G. Ellenbogen, J.N. Oshinski, B.J. Iskandar, and B.A. Martin, "Cerebellar tonsil ectopia measurement in type I Chiari malformation patients show poor inter-operator reliability," *Fluids Barriers CNS.*, vol. 15, no. 1, pp. 33, 2018.

[7] J.W. McVige, and J. Leonardo, "Neuroimaging and the clinical manifestations of Chiari Malformation Type I (CMI)," *Curr Pain Headache Rep.*, vol. 19, no. 6, pp. 18, 2015.

[8] R.F. Jr. Sekula, P. Jannetta, K.F. Casey, E.M. Marchan, L.K. Sekula, and C.S. McCrady, "Dimension of the posterior fossa in patients symptomatic for Chiari I malformation but without cerebellar tonsillar descent," *Cerebrospinal Fluid research*, vol. 2, pp. 11–18, 2005.

[9] J. Meadows, M. Kraut, and M. Guarnieri, "Asymptomatic Chiari type I malformations identified on magnetic resonance imaging," *J Neurosurg.*, vol. 92, pp. 920–6, 2000.

[10] C. Mottolese, A. Szathmari, and E. Simon, "Treatment of Chiari type I malformation in children: the experience of Lyon," *Neurolog Sci.*, vol. 32 (suppl 3), pp. S325–30, 2011.

[11] R.S. Tubbs, S. Elton, and P. Grabb, "Analysis of the posterior fossa in children with the Chari 0 malformation," *Neurosurgery*, vol. 48(5), pp. 1050–5, 2001.

[12] E.N. Kahn, K.M. Muraszko, and C.O. Maher, "Prevalence of Chiari I malformation and syringomyelia," *Neurosurg Clin N Am.*, vol. 26, pp. 501–507, 2015.

[13] J. Meadows, M. Kraut, M. Guarnieri, R.I. Haroun, and B.S. Carson, "Asymptomatic Chiari type I malformations identified on magnetic resonance imaging," *J Neurosurg.*, vol. 92, pp. 920–926, 2000.

[14] T.H. Milhorat, M. Nishikawa, R.W. Kula, and Y.D. Dlugacz, "Mechanisms of cerebellar tonsils herniation in patients with Chiari malformation as guide to clinical management," *Acta Neurochir.*, vol. 152, pp. 1117–1127, 2010.

[15] Y.W. Wu , C.T. Chin, K.M. Chan, A.J. Barkovich, and D.M. Ferriero, "Pediatric Chiari I malformations: do clinical and radiologic features correlate?" *Neurology*, vol. 53, pp. 1271–1276, 1999.

[16] J. Siasios, E.Z. Kapsalaki, and K.N. Fountas, "Surgical management of patients with Chiari I malformation," *Int J Ped.*, ID 640127, 2012.

[17] J. Baisden, "Controversies in Chiari I malformations," *Surg Neurol Int.*, vol. 3 (Suppl 3), pp. S232–S237, 2012.

[18] R.S.Tubbs, J. Beckman, and R.P. Naftel, "Istitutional experience with 500 cases of surgically treated pediatric Chiari malformation type I," *J Neurosurg Ped.*, vol. 7, pp. 248–56, 2011.

[19] L. Genitori, P. Peretta, and C. Nurisso, "Chiari type I anomalies in children and adolescents: minimally invasive management in a series of 53 cases," *Childs Nerv Syst.*, vol. 16, pp. 707–18, 2000.

[20] M. Caldarelli, F. Novegno, and L. Vassimi, "The role of limited posterior fossa craniectomy in the surgical treatment of Chiari malformation type I: experience with a pediatric series," *J Neurosurg.*, vol. 106, pp. 187–95, 2007.

[21] B.C. Kennedy, K.M. Kelly, and M.Q. Phan, "Outcomes after suboccipital decompression without dural opening in children with Chiari malformation type I," *J Neurosurg Ped.*, vol. 16, pp. 150–8, 2015.

[22] F.M. Limonadi, and N.R. Selden, "Dura-splitting of the craniocervical junction: reduced operative time, hospital stay, and cost with equivalent early outcome," *J Neurosurg.*, vol. 101 (2 suppl), pp. 184–8, 2004.

[23] S. Kalb, L. Perez-Orribo, M. Mahan, N. Theodore, P. Nakaji, and R.E. Bristol, "Evaluation of operative procedures for symptomatic outcome after decompression surgery for Chiari type I malformation," *J Clin Neur.*, vol. 19, pp. 1268–1272, 2012.

[24] E. Schijman, and P. Steinbock, "International survey on the management of Chiari I malformation and syringomyelia," *Child Nerv Syst.*, vol. 20, pp. 341–348, 2004.
[25] L.A, Aitken, C.E. Lindan, and S. Sidney, "Chiari type I malformation in a pediatric population," *Pediatr Neurol.*, vol. 40(6), pp. 449–54, 2009.
[26] A. Urbizu, M.A. Poca, X. Vidal, A. Rovira, J. Sahuquillo, and A. Macaya, "MRI-based morphometric analysis of posterior cranial fossa in the diagnosis of Chiari malformation type I," *J Neuroimaging*, vol. 24, no. 3, pp. 250–256, 2014.
[27] A.M. Bagci, S.H. Lee, and N. Nagornaya, "Automated posterior cranial fossa volumetry by MRI: applications to Chiari malformation type I," *AJNR Am J Neuroradiol.*, vol. 34 (9), pp. 1758–63, 2013.
[28] S.V. Furtado, K. Reddy, and A.S. Hedge, "Posterior fossa morphometry in symptomatic pediatric and adult Chiari I malformation," *J Clin Neurosci.*, vol. 16, pp. 1449–54, 2009.
[29] M. Nishikawa, H. Sakamoto, and A. Hakuba, "Pathogenesis of Chiari malformation: a morphometric study of the posterior cranial fossa," *J Neurosurg.*, vol. 86, pp. 40–7, 1997.
[30] T. Trigylidas, B. Baronia, and M. Vassilyadi, "Posterior fossa dimension and volume estimates in pediatric patients with Chiari I malformation," *Childs Nerv Syst.*, vol. 24, pp. 329–36, 2008.
[31] I. Despotovic, B. Goossens, and W. Philips, "MRI Segmentation of the Human Brain: Challenges, Methods, and Applications," Comput Math Methods Med. vol. 2015, 450341, 2015.
[32] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, and N. Delcroix, "Automated anatomical labelling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single subject brain," Neuroimage vol. 15, pp. 273–89, 2002.
[33] A. Keszei, B. Berkels, and T. Deserno, "Survey of non-rigid registration tools in medicine," J Digit Imaging, vol. 32, no. 30, pp. 102–16, 2017.
[34] T. Fu, Q. Li, J. Zhu, D. Ai, Y. Huang, H. Song, Y. Jiang, Y. Wang, and J. Yang, "Sparse deformation prediction using markove decision processes (mdp) for non-rigid registration of mr image," Comput Methods Programs Biomed, vol. 162, pp. 47–59, 2018.
[35] L. Fang, L. Zhang, D. Nie, X. Cao, I. Rekik, S. Lee, H. He, and D. Shen, "Automatic brain labeling via multi-atlas guided fully convolutional networks," Med Image Anal, vol. 51, pp. 157–68, 2019.
[36] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G.Z. Yang, "Deep Learning for Health Informatics," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 4–21, 2017.
[37] J.-P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243-260, 1998.
[38] O.V. Senyukova, and A.Y. Zubov, "Full anatomical labeling of magnetic resonance images of human brain by registration with multiple atlases," Program Comput Soft, vol. 42, pp. 356–60, 2016.
[39] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Intern J of Comput Vis.*, pp. 321–331, 1988.
[40] Y. Wei, J. Zhou, Y. Wang, Y. Liu, Q. Liu, J. Luo, C. Wang, F. Ren, and L. Huang "A Review of Algorithm & Hardware Design for AI-Based Biomedical Applications," IEEE Transactions on Biomedical Circuits and Systems, vol. 14, no. 2, pp. 145–63, 2020.
[41] L. Mesin, F. Mokabberi, and C.F. Carlino, "Identification of optimal surgical intervention for Chiari I malformation," 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Siena, Italy, 2019.
[42] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. 61–72, 2009.
[43] L. Mesin, "Balanced multi-image demons for non-rigid registration," Magnetic Resonance Imaging, submitted, 2019.
[44] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: a survey," IEEE Trans Med Imaging, vol. 32, no. 7, 1153–90, 2013.
[45] K. Zuiderveld, "Contrast Limited Adaptive Histograph Equalization," Graphic Gems IV. San Diego: Academic Press Professional, pp 474–485, 1994.
[46] P. Kovesi, "Symmetry and Asymmetry from Local Phase," Tenth Australian Joint Converence on Artificial Intelligence, pp. 2–4, 1997.
[47] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Symmetric log-demons: diffeomorphic image registration," International Journal of Computer Vision, vol. 107, no. 3, pp. 254–271, 2014.
[48] H. Lombaert, L. Grady, X. Pennec, N. Ayache, and F.Cheriet, "Spectral Log-Demons: Diffeomorphic Image Registration with Very Large Deformations." International Journal of Computer Vision, Springer Verlag, vol. 107, no. 3, pp. 254–71, 2014.
[49] H. Trevor, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York Inc., USA, 2001.
[50] I. Tsamardinos, A. Rakhshani, and V. Lagani, "Performance-estimation properties of cross-validation-based protocols with simultaneous hyper-parameter optimization," in Proc. 8th Hellenic Conf. Artif. Intell.: Methods Appl., pp. 1–14, 2014.
[51] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Vol. 9351 of LNCS, Springer, pp. 234–241, 2015.
[52] J. Peyrat, H. Delingette, M. Sermesant, C. Xu, and N. Ayache, "Registration of 4D cardiac ct sequences under trajectory constraints with multichannel diffeomorphic demons," IEEE Trans Med Imag., vol. 29, no. 7, pp. 1351–68, 2010.
[53] H. Lombaert, L. Grady, X. Pennec, J.-M. Peyrat, N. Ayache, and F. Cheriet, "Groupwise spectral log-demons framework for atlas construction," Medical Computer Vision (MCV'12) MICCAI workshop, 2012.
[54] W. Huizinga, D. Poot, J. Guyader, R. Klaassen, B. Coolen, M. van Kranenburg, R. van Geuns, A. Uitterdijk, M. Polfliet, J. Vandemeulebroucke, A. Leemans, W. Niessen, and S. Klein, "PCA-based groupwise image registration for quantitative MRI," Med Image Anal., vol. 29, pp. 65–78, 2016.

Luca Mesin graduated in Electronics Engineering in 1999 and received the Ph.D. in Applied Mathematics in 2003 from Politecnico di Torino, Italy. From 2003 to 2008 he was a Fellow of the Laboratory for Neuromuscular System Engineering of the Department of Electronics, Politecnico di Torino. Since 2008, he is Assistant Professor in Biomedical Engineering at the Department of Electronics and Telecommunications and head of the Mathematical Biology and Physiology group (Politecnico di Torino). His main research activities are in the fields of biomedical image/signal processing and mathematical modelling.

Forough Mokabberi graduated in Electronics Engineering in 2014 from Guilan University, Iran, and received the Master of Science in Biomedical Engineering in 2018 from Politecnico di Torino, Italy. Her main research activities are in the fields of medical image processing and data classification.

Christian Francesco Carlino graduated in Medicine and Surgery in 2001 and obtained his specialization in Neurosurgery in 2006, at the University of Turin. From 2010 to 2017 he worked as a child neurosurgeon at the Regina Margherita Children's Hospital in Turin. From 2017 he works at the Neurosurgery Department of San Giovanni Bosco Hospital in Turin. His main field of interest concerns malformative pathologies of the central nervous system and disorders of the state of consciousness.