An Automatic Multiresolution Atlas-based Framework for Segmentating Tibia, Femur and Knee Cartilage on MRI Volumes

José Bernal, Èric Pairet and Songyou Peng
University of Girona

***Abstract*— Medical images are used in medicine to diagnose a wide range of health illnesses. Specially, this source of information takes special importance when it comes to one of the most frequently injured parts of the body, the knee. To make this images more understandable to doctors, the state-of-the-art proposes many techniques to segment the different structural elements of this joint, e.g. tibia, femur and the corresponding knee cartilages. In this paper, a segmentation framework has been proposed to segment tibia, femur and its corresponding knee cartilages on Magnetic Resonance Imaging (MRI) volumes. Specifically, the proposal first builds an atlas and an anatomical model for each desired structure of the knee, which are used later in a Bayesian classification framework. The obtained output is then introduced as a set of seeds in a geodesic active contours method to obtain the final segmentation. The results have shown that the proposal has a huge potential for segmenting the different structures of the knee, which could be dramatically improved by taking into account the proposed future works.**

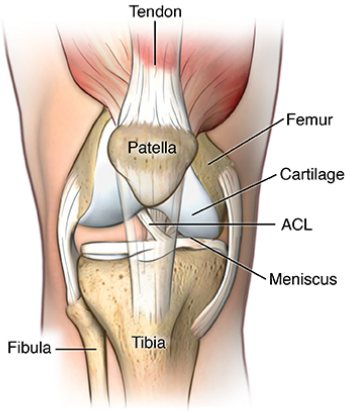


Fig. 1: Parts of the knee. Image taken from: http://www.hopkinsmedicine.org.

I. INTRODUCTION

Nowadays, one of the main challenges of medicine and engineering is to develop tools for supporting medical diagnosis, treatment and surgeries. In particular, knee-related procedures are of interest since this complex joint is one of the most frequently injured parts in our body. According to the World Health Organization (WHO), chronic rheumatic conditions, such as osteoarthritis, form part of the top ten causes of disability in developed countries [1]. Moreover, anterior cruciate ligament injuries – related to the damage on the surface of femur and tibia – determines the retirement of a high percentage of the athletes [2]. Thus, detecting problems in the knees correctly and timely may improve the life quality of a considerable portion of the world population.

The knee is a joint formed by different anatomical structures as presented in Fig. 1. Femur (thigh bone) and tibia (shin bone) are two major bones of knees with soft tissues like cartilage covering their connected regions. In this paper, we address the segmentation of tibia, femur and cartilage around them.

Several algorithms can be found in the literature for performing segmentation. Promising results have been reached in medical imaging by applying algorithms such as Otsu's thresholding [3], region growing [4], deformable models [5], [6] and atlas-guided approach [7]. Notably, altas-guided approach takes into account spatial and shape information to address this task. In this report, a probabilistic atlas-based approach for segmenting 3D MRI volumes using a Bayesian framework is presented.

The report is organised as follows. Some state-of-the-art knee segmentation methods are discussed in Section II. Our proposed method are detailed in Section III. The method was evaluated with a set of volumes and qualitative and quantitative results are presented in Section IV. The discussion and analysis of the obtained results is exposed in Section V Finally, remarks and proposals of future work are given in Section VI and Section VII.

II. STATE-OF-THE-ART

In the last years, a large number of methods have been applied for segmenting knee structures. For instance, Kapur *et al.* [8] proposed an adaptive region growing method which analyses texture information of different seeded regions; Folkesson *et al.* [9] segmented cartilage using a voxel-based hierarchical classification scheme; the method of Yin *et al.* [10] simultaneously segmented multiple objects based on a layered optimal graph.

Different deformable models, such as Active Contour Model (ASM), Active Appearance Models (AAM) and probabilistic models are among the state-of-the-art methods which are demonstrated to provide promising segmentation results. The ASM has been used by Solloway *et al.* [11] to automatically segment femoral cartilages using B-spline interpolation. Also, it has been used by Fripp *et al.* [12] for femur and tibia segmentation and cartilage extraction. The last task has been also addressed by Vincent *et al.* [13] using a hierarchical AAM technique.

Multi-atlas segmentation strategies have already been successfully applied in brain imaging segmentation [14] due

to its advantage of providing the spatial information and suppressing unnecessary interference from other tissues or organs. However, these methods have been rarely used in knee segmentation. In this report, we used the atlas-based segmentation for the main parts of the knee, e.g. femur, tibia and the corresponding cartilages. In particular, the main contributions of our work are:

1) A multiresolution affine registration framework to accelerate the construction of multiple atlases.
2) A region growing method based on geodesic active contour [15] to improve the initial segmentation results by giving a better boundary fitting.

III. PROPOSED APPROACH

As mentioned previously, the aim of the segmentation algorithm proposed in this paper is to classify tibia, femur and cartilage from MRI volumes. The overall segmentation framework is based on the approaches of Park *et al.* [7] and Gubern-Mérida *et al.* [16] but many modifications were introduced such that it is feasible for our specific case.

The segmentation framework consists mainly of the two processes illustrated in Fig. 2: (i) training process in which probabilistic atlases and anatomical models are defined and (ii) segmentation process of a given volume. These two procedures are subdivided into different tasks which are detailed in the following sections.

A. Image pre-processing

The given dataset may contain images with uninteresting variability caused by acquisition conditions. Even when considering the same patient and the same device, the gathered images will not be the same. Thus, pre-processing techniques are considered to enhance the quality and decrease the distortions before computing the atlases and anatomical models. The considered methods are normalisation given by bias field correction and contrast stretching, mirroring of volumes, re-sampling and padding. The details of these operations are presented below.

1) Normalization: The images in the dataset may have the next problems: the intensity of voxels of the same tissue may vary in the same image (inter-patient variability) and, also, among patients (intra-patient variability). In Fig. 3, two volumes of the same part of the body evidencing the latter problem are presented. In this case, the range of the intensities of the first image is $[0, 400]$ while in the second is $[0, 3500]$. Then, the goal of the normalisation step is to reduce these alterations.

In this step, we address the normalisation problem two-fold: (i) improved non-parametric bias field correction (N4) [17], [18] to deal with inter-patient variance and (ii) contrast stretching correction [19] to reduce intra-patient variability.

2) Mirroring: It is well-known that bones of left and right legs evidence different shapes as shown in Fig. 4. If this fact is not considered, the obtained atlas will not be accurate. For the convenience of the model, we assume that one could be transformed into the other using a mirroring transformation

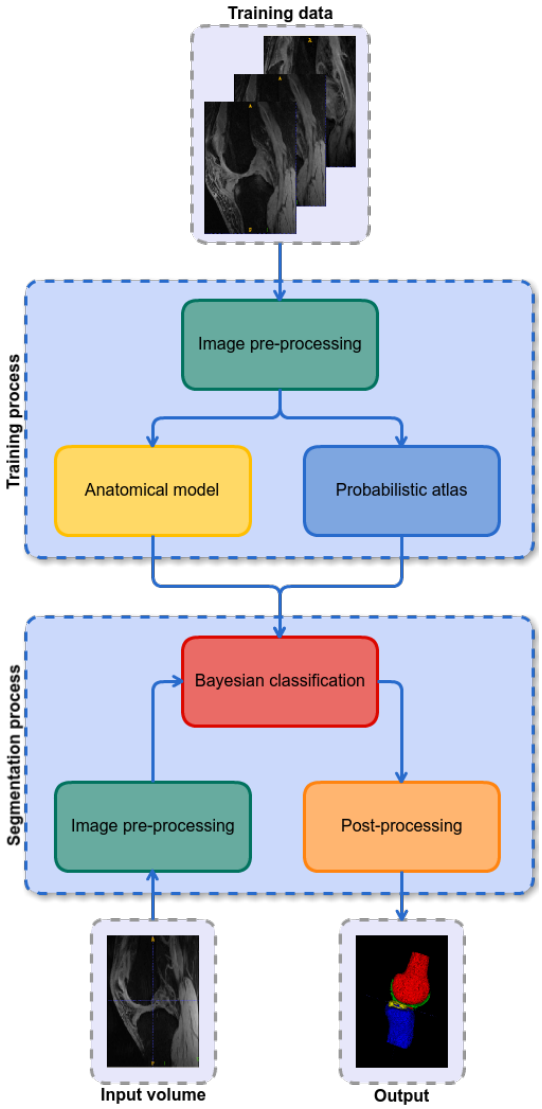


Fig. 2: Proposed segmentation framework.

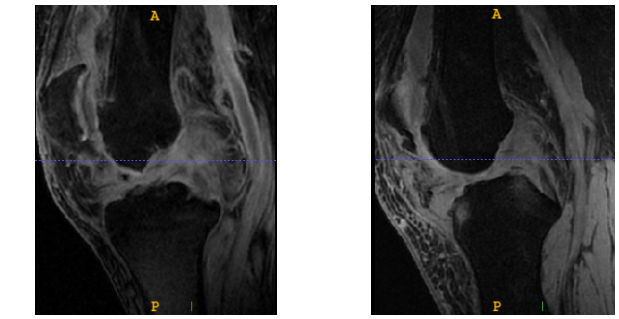
along the z-axis. The images are mirrored and transformed back if necessary after obtaining final segmentation.

3) Re-sampling: Recall that an image is actually the sampling of the continuous field using a discrete grid. The physical distance between neighbouring points in the grid is called spacing. This factor, determined in part by characteristics of the acquisition device, may be not the same for all the images. Thus, another pre-processing step called re-sampling is used to unify the spacing of all the inputs; in this particular case, the spacing was fixed to $0.390625 \times 0.390625 \times 1$.

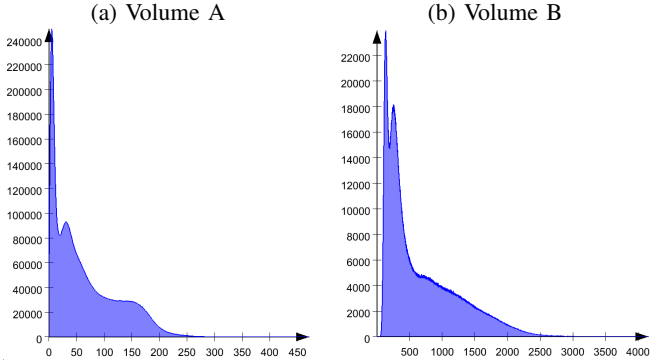
4) Padding: After re-sampling all the images, we may end up having volumes of different dimensions. Thus, the approach consists in finding the maximum height, width and depth in the database to later pad them all with zeros to this maximum size.

B. Construction of probabilistic atlas

After pre-processing all the training images, the probabilistic atlases for femur, tibia, cartilages and background are

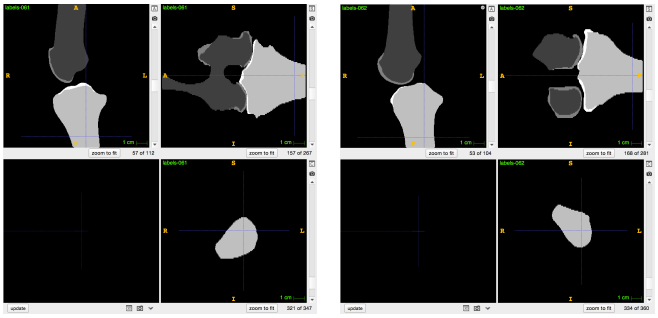


(a) Volume A (b) Volume B



(c) Histogram of Volume A (d) Histogram of Volume B

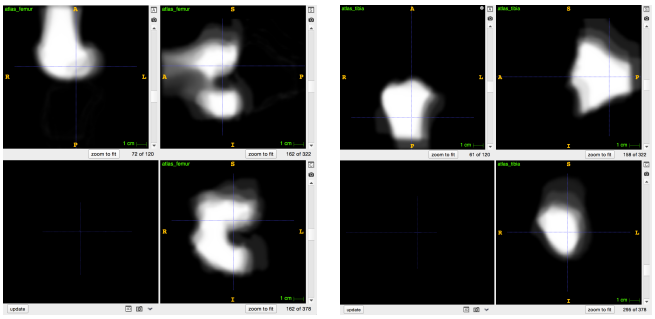
Fig. 3: Intensity variability. Both images correspond to MRI of knees, but the range of intensities differs between them.



(a) Left leg (b) Right leg

Fig. 4: Visual difference between left and right legs.

calculated. Firstly, a reference image from the training set is chosen. This image should be carefully selected since it may dramatically affect the final output. Secondly, each remaining image is registered with respect to the reference. This process is carried out by using the label images instead of the MRI. Thirdly, the obtained transformation is applied to the original image. This approach has been considered due to the fact that it is easier and faster to obtain the transformations when less data is involved. Moreover, even after the normalisation process, the tissues around the bones may complicate the registration process. Finally, the probabilistic atlases are calculated by averaging the different regions of interest separately. Note that the acquired atlases will evidence different layers as a consequence of its calculation. Since this aspect may limit the area belonging to a certain class, a Gaussian filter with a standard deviation of 1.5 is considered. To this



(a) Atlas of femur (b) Atlas of tibia

Fig. 5: Atlases of femur and tibia after registration and smoothing.

point, the atlases evidence a smooth decrease of intensities and uncertainty around the boundaries.

Note that the domain of the atlas is within the range $[0, 1]$, which indicates the probability of belonging or not to a specific class. Also, after acquiring the atlas for organs and tissues, the atlas for the background can be simply represented as an image like $(1 - \text{organs} \cup \text{tissues})$. Some resulting atlases are presented in Fig. 5.

The computational cost of the image registration framework depends on the size of the volumes that are processed. The bigger the image, the more the algorithm will take to process it. Since the aim of this registration process is to obtain results in a short period of time, normal approaches may not be suitable. However, it can be speeded up by embedding it into a multiresolution framework.

The idea of multiresolution is to build image pyramids for the fixed and moving images in which each level corresponds to a scaled version of the originals as presented in Fig. 6. The scale decreases from bottom to top, being the initial image at the bottom and the smallest image at the top of the pyramid. The process behind multiresolution is the following: (1) the fixed and moving images at the top of the pyramid are registered, (2) the obtained result is scaled according to the scaled difference between two layers of the pyramid, (3) the scaled parameters are taken into account for initializing the next level. The process is carried out until the full pyramid is explored.

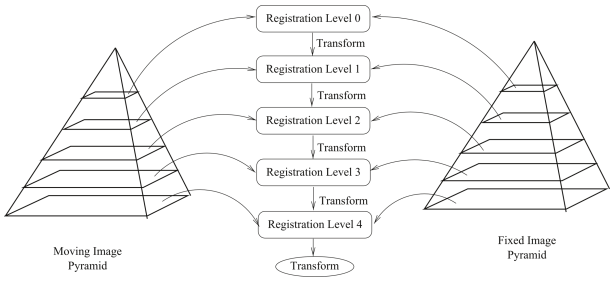


Fig. 6: Multiresolution framework. Image taken from: http://yadics.univ-lille1.fr.

The considered transformation embedded in the registration process corresponds to an affine transformation which

takes into account scaling, shearing, translation and rotation. Note that we do not consider non-rigid transformation since tibia and femur are not subjected to deformations as soft tissues.

C. Construction of the anatomical model

As the atlas contains information regarding location, distribution and shape of the object that it represents, the anatomical model incorporates knowledge about the range of intensities. Note that this process can be carried out since inter- and intra-patient normalisations were performed previously.

The construction of the anatomical model is the following. Firstly, the histogram of each region of interest is computed. The result would be similar to the one presented in Fig. 7. Secondly, we assume that each class – except for the class 'none' – can be represented by a Gaussian. Hence, the mean μ_k and standard deviation σ_k are calculated for each one of them.

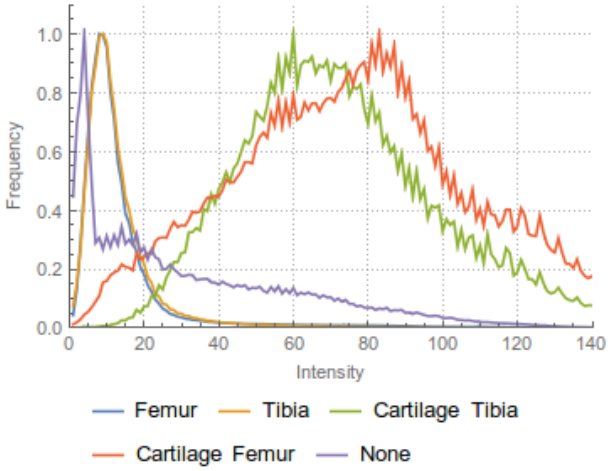


Fig. 7: Histograms of intensities for different types of tissues or organs according to the information in the dataset.

The class 'none' represents an exception since it may involve different tissues or organs with varied intensities and, hence, one mean and one standard deviation would not be enough for properly representing the data. Instead, a multimodal distribution represented by a mixture of Gaussians is assumed and, therefore, a group of parameters are estimated.

The calculated parameters are used for determining the probability that a voxel belongs to a certain class given that the intensities of the class are within the obtained ranges. Given a pre-processed test image Y formed by voxels y_i and the set of classes X, such probability is expressed as follows:

$$p(y_i \in Y|\mathbf{X}_i = k) = \frac{1}{\sqrt{2\pi\sigma_k^2}}\exp\left(\frac{-(y_i-\mu_k)^2}{2\sigma_k^2}\right), \quad k = 0,1,2,3,4; \tag{1}$$

where $k = 0,1,2,3,4$ represent background, femur, its surrounding cartilage, tibia and its surrounding cartilage, respectively. Note that the higher the probability, the more we can be sure that the voxel actually is related to the class. However, this information should be used along with the atlas to get the final verdict.

D. Bayesian segmentation framework

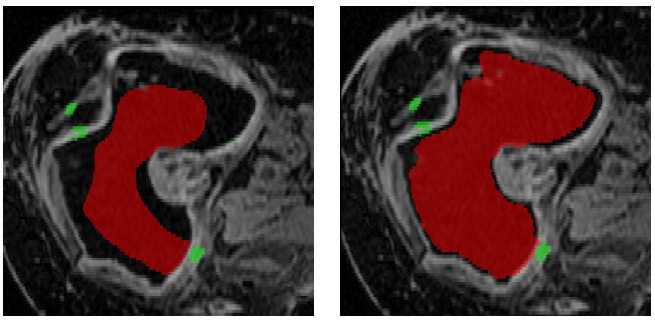
Once the probabilistic atlases of training images and the tissue models for the input image are acquired, segmentation can be performed. Image segmentation aims to estimate the label $\mathbf{X}$ that can best describe the given testing image $\mathbf{Y}$, for which the simple but efficient Maximum A Posteriori (MAP) is quite suitable. This can be modelled by the following expression:

$$P(\mathbf{X}|\mathbf{Y}) \propto \arg\max P(\mathbf{X})P(\mathbf{Y}|\mathbf{X}), \tag{2}$$

where the prior $P(\mathbf{X})$ is given by the probabilistic atlas and the likelihood $P(\mathbf{Y}|\mathbf{X})$ is the intensity tissue models calculated in Sec. III-C.

E. Post-processing

After the initial segmentation from Sec. III-D, we acquire the result shown in the Fig. 8a. As expected, the segmentation is within the region of interest due to the fact that the atlases for femur and tibia with the values over 0.96 are applied in the Bayesian framework. However, the drawback is that the initial segmentation does not cover the entire region of interest. Thus, we need to perform post-processing to enhance the result. The approach is described as follows. First of all, the pre-processed input volume should be resized back. Additionally, if the image was mirrored in the pre-processing step, it is necessary to transform it back. Afterwards, the boundaries may not be as defined as expected. One way to deal with it is to consider high-boost. In this case, we consider a recursive Laplacian filter for extracting the edges. Moreover, since negative values may exist on the boundaries after applying the last operation, the absolute value is taken into account. After that, histogram equalisation is performed in order to improve the contrast among different organs and tissues. Finally, the segmentation results given by the Bayesian classification are set as seeds, which are later used by the geodesic active contour level set filter. The initial structure grows as the boundaries of the ROI allow it to. (Fig. 8b).



(a) Before post-processing (b) After post-processing

Fig. 8: Illustrations for the post-processing.

F. Expected limitations

After extensive observations, three limitations were recognised. First of all, enough data is required for building atlases as well as anatomical models. If only a few volumes are considered for training, the spatial information represented by the atlases and the parameters used for estimating the Gaussians may be not representative. Additionally, if the bias field correction is not performed correctly, the bias of the intensity values inside a certain organ or tissue may remain large and, hence, leakages may appear in the final segmentation results. Moreover, if the boundaries are not clearly separable, leakage will still happen even though bias field is performed perfectly. And, last but not least, the performance of registration is an essential part of the process. If the registration is not properly carried out, the assumptions contained in $P(X)$ might be wrong since the atlases are placed in incorrect positions.

IV. RESULTS

In this section, the segmentation results obtained using the proposed approach on a subset of the SKI10 MRI dataset were evaluated using Dice Similarity Coefficient (DSC), sensitivity, specificity and computing time.

The DSC measure is a ratio between the intersection and the union of two sets. As they get closer, the higher its value and, hence, the better the classification. The results of the segmentation using DSC are presented in Fig. 9. On one hand, it can be observed that femur and tibia obtained scores above 0.8 in the majority of the cases. This may indicate that the framework was able to place the seeds in the correct position but the post-processing had some leakages caused by a low contrast between bones and other surrounding tissues as a result of the bias field. On the other hand, the classification for cartilage of femur and tibia was not able to surpass 0.2. In general, femur was the best-segmented class compared to the others.

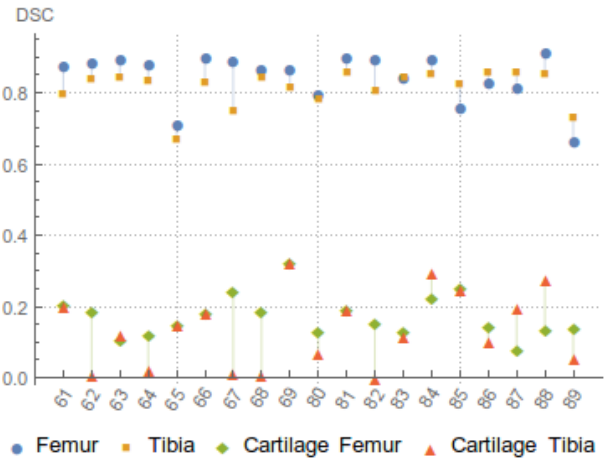


Fig. 9: Dice Similarity Coefficient per region of interest.

The sensitivity of an algorithm determines its ability to identify positive results. Thus, the higher its value, the more discriminative the proposal is between normal or abnormal samples. The results of this measure are presented in Fig. 10. It can be observed that the proposed approach can classify femur and tibia with a sensitivity of 0.8 in most of the cases. However, the results for cartilages do not surpass 0.2.

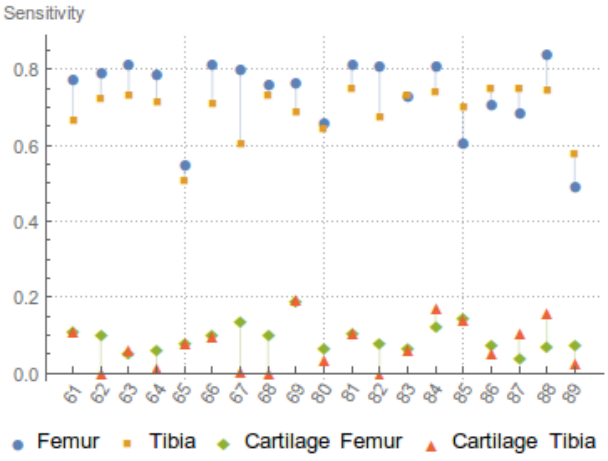


Fig. 10: Sensitivity per region of interest.

The specificity of an algorithm determines its ability to identify negative cases. Thus, the higher the specificity, the less number of incorrectly classified voxels during the segmentation. The results using this measure are presented in Fig. 11. It can be seen that the specificity is higher than 0.95 for all the situations, i.e. at the 95% of the cases, the algorithm spotted correctly the negative cases of each class. However, this result is severely affected by the difference in size of the background with respect to the other classes.

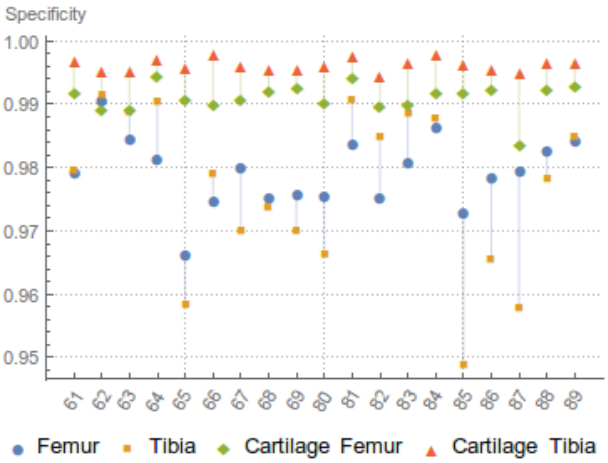


Fig. 11: Specificity per region of interest.

The SKI10 evaluation framework considers some additional measures and constraints to determine the score of a segmentation. In the case of bones, the average symmetric surface distance and the root-mean-square symmetric surface distance was used to compute the score. In the case of cartilage, the volume and thickness of the segmentation

determined by the volumetric overlap error and the volumetric difference are considered. The final results under this assessment scheme are presented in Fig. 12. It can be seen that the proposal was able to achieve better results for cartilage than for bones. This outcome may be a consequence of leakages in the post-processing step.

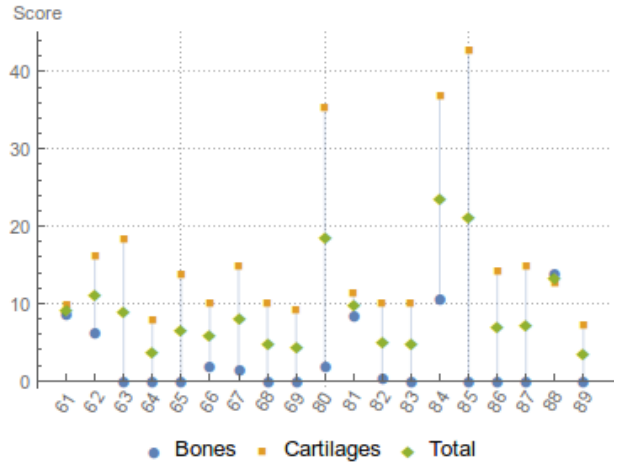


Fig. 12: Score for bones and cartilages according to SKI10 evaluation framework.

The entire procedure from pre-processing the images to the final segmentation takes around 11 minutes per input image using the multiresolution approach with 3 levels and 2 iterations per layer. This means that the proposed segmentation framework is suitable for application in which response in short periods of time is required.

V. DISCUSSION

As observed in the last section, the evaluation highlighted some difficulties of the approach for correctly identifying cartilages. Four situations might be generating this issue:

- As mentioned previously, the registration step is crucial since it determines the transformation to apply to the atlases to classify the input image. Thus, if the registration was not good enough, the atlases for cartilages were placed in locations where there was no match for these classes. However, since the same matrix was used to transform the different atlases and the segmentation for bones achieved better results, it can be concluded that the registration was performed correctly.
- A bad registration during the training process may result in a bad atlas. In this case, the atlas would not be able to condense information of cartilages. This can be also seen if the samples provided to the framework were not representative.
- The anatomical model was not able to describe the range of intensities in which the cartilage values were moving, i.e. the assumption that a Gaussian distribution represented the intensities of cartilages was incorrect. In order to improve this section, a further analysis should be performed to determine the convenience of the normal distribution in this case.
- The size of the tissue in comparison to the other areas of interest made the process difficult. As mentioned before, a Gaussian smoothing kernel is applied over the atlases in order to add uncertainty to the bordering areas. In the case of background, the areas around the bones are close to 1 even after filtering and, hence, it may decrease the chances of classifying the cartilage correctly. One way to improve this issue is to weight the tissue over the background.

At the same time, the SKI10 evaluation highlighted some issues on the segmentation of bones. Two limitations of the approach may drive to this situation. On one hand, high presence of bias field on a tissue results in a large variation of intensities within an area of interest. Thus, in the post-processing step, the initial seeds are not expanded to these locations with different grey levels. On the other hand, if there is no contrast between two different areas of interest, leakages may occur and, hence, a larger difference between the expected volume and the obtained one takes place. This situation may be improved by considering higher bias field correction or contrast-enhancing algorithms.

VI. FINAL REMARKS

An automatic multiresolution atlas-based framework for segmenting tibia, femur and the cartilages surrounding them was proposed, implemented and evaluated under the scheme of the SKI10 challenge.

In short, the proposal considers gathering atlases and anatomical models for the different classes from a set of MRI volumes. This information is later used in a Bayesian classification framework for processing a given image. Then, the obtained result is used as seeds for a region growing method based on geodesic active contours in order to obtain the final segmentation.

The results have shown that the approach was able to classify tibia and femur with high accuracy compared to cartilage. However, the score using the evaluation framework reflected the opposite since it considers some additional constraints.

Some drawbacks of the proposal were analysed during the evaluation of the results. It has been highlighted that a good registration, the assumption of the correct distributions to describe the anatomical models and, also, a significant reduction of the bias field are fundamental in order to obtain accurate segmentations.

VII. FUTURE WORK

We recommend evaluating different ways to improve the contrast between bones and surrounding tissues such that the leakage can be avoided and, therefore, the segmentation results can be improved.

Another future improvement of the presented proposal is to consider the real distribution of intensities of the different classes rather than assuming Gaussians or mixtures of them.

Finally, we encourage the readers to consider a Markov Random Field to regularise the final segmentation results.

REFERENCES

[1] World Health Organization. Chronic diseases and health promotion. http://www.who.int/chp/topics/rheumatic/en/. [Online; accessed 2016-05-05].

[2] Jeffrey Kay, Jon Karlsson, Volker Musahl, Olufemi R Ayeni, et al. Anterior cruciate ligament rupture a family affair. *Orthopaedic journal of sports medicine*, 3(11):2325967115616783, 2015.

[3] Liu Jianzhuang, Li Wenqing, and Tian Yupeng. Automatic thresholding of gray-level pictures using two-dimension otsu method. In *Circuits and Systems, 1991. Conference Proceedings, China., 1991 International Conference on*, pages 325–327 vol.1, Jun 1991.

[4] Stephen Gould, Tianshi Gao, and Daphne Koller. Region-based segmentation and object detection. In *Advances in neural information processing systems*, pages 655–663, 2009.

[5] Bram Van Ginneken, Alejandro F Frangi, Joes J Staal, Bart M Romeny, and Max A Viergever. Active shape model segmentation with optimal features. *Medical Imaging, IEEE Transactions on*, 21(8):924–933, 2002.

[6] Gareth J Edwards, Christopher J Taylor, and Timothy F Cootes. Interpreting face images using active appearance models. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 300–305. IEEE, 1998.

[7] Hyunjin Park, Peyton H Bland, and Charles R Meyer. Construction of an abdominal probabilistic atlas and its application in segmentation. *Medical Imaging, IEEE Transactions on*, 22(4):483–492, 2003.

[8] Tina Kapur, P Beardsley, S Gibson, W Grimson, and W Wells. Model-based segmentation of clinical knee mri. In *Proc. IEEE Intl Workshop on Model-Based 3D Image Analysis*, pages 97–106. Citeseer, 1998.

[9] Jenny Folkesson, Erik B Dam, Ole F Olsen, Paola C Pettersen, and Claus Christiansen. Segmenting articular cartilage automatically using a voxel classification approach. *Medical Imaging, IEEE Transactions on*, 26(1):106–115, 2007.

[10] Yin Yin, Xiangmin Zhang, Rachel Williams, Xiaodong Wu, Donald D Anderson, and Milan Sonka. Logismoslayered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint. *Medical Imaging, IEEE Transactions on*, 29(12):2023–2037, 2010.

[11] Stuart Solloway, Chris J Taylor, Charles E Hutchinson, and John C Waterton. Quantification of articular cartilage from mr images using active shape models. In *Computer VisionECCV'96*, pages 400–411. Springer, 1996.

[12] Jurgen Fripp, Stuart Crozier, Simon K Warfield, and Sébastien Ourselin. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. *Medical Imaging, IEEE Transactions on*, 29(1):55–64, 2010.

[13] Graham Vincent, Chris Wolstenholme, Ian Scott, and Mike Bowes. Fully automatic segmentation of the knee joint using active appearance models. *Medical Image Analysis for the Clinic: A Grand Challenge*, pages 224–230, 2010.

[14] Torsten Rohlfing, Robert Brandt, Randolf Menzel, and Calvin R Maurer. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4):1428–1442, 2004.

[15] Michael E Leventon, W Eric L Grimson, and Olivier Faugeras. Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 316–323. IEEE, 2000.

[16] Albert Gubern-Mérida, Michiel Kallenberg, Robert Martí, and Nico Karssemeijer. Multi-class probabilistic atlas-based segmentation method in breast mri. In *Pattern Recognition and Image Analysis*, pages 660–667. Springer, 2011.

[17] John G Sled, Alex P Zijdenbos, and Alan C Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *Medical Imaging, IEEE Transactions on*, 17(1):87–97, 1998.

[18] Nicholas J Tustison, Brian B Avants, Philip A Cook, Yuanjie Zheng, Alexander Egan, Paul A Yushkevich, and James C Gee. N4itk: improved n3 bias correction. *Medical Imaging, IEEE Transactions on*, 29(6):1310–1320, 2010.

[19] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.