IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 27, 2020, accepted February 21, 2020, date of publication March 19, 2020, date of current version March 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982061

Personalized Knee Geometry Modeling Based on Multi-Atlas Segmentation and Mesh Refinement

FILIPPOS P. NIKOLOPOULOS, EVANGELIA I. ZACHARAKI, DIMITAR STANEV, AND KONSTANTINOS MOUSTAKAS, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, University of Patras, 26504 Patras, Greece

Corresponding author: Filippos P. Nikolopoulos (filnikolo@gmail.com)

This work was supported in part by the EC Horizon 2020 project OActive: Advanced personalized, multi-scale computer models preventing osteoarthritis, under Grant 777159.

ABSTRACT The development of personalized finite element models of the knee anatomy is critically important in the simulation of knee joint mechanics, prediction of optimal treatments in cases of pathological conditions and prevention of injuries. Subject-specific models can be obtained from diagnostic images with multi-atlas segmentation being a pertinent choice when prior anatomical information of the structures of interest is available. Although multi-atlas segmentation has been prevalent in some parts of the body, its exploitation for the segmentation of the knee complex has not been illustrated yet. This work utilizes a multi-atlas segmentation method based on deformable registration and joint label fusion in conjunction with anatomically-adopted mesh refinement in order to generate subject-specific models of the knee. The success of finite element simulations strongly depends on the properties of the 3D surface and the quality of the volumetric meshes. Therefore, emphasis was given to create structured meshes with well-shaped hexahedra for the knee cartilages and menisci. The segmentation performance is assessed using cross-validation on 7 subjects from the Open Knee project and 78 subjects from the Osteoarthritis Initiative. Our results indicate that our developed state-of-the-art processing scheme can achieve competitive performance, opening the path for better diagnostics and patient-specific interventions. The developed tools are freely available to download from SimTK at https://simtk.org/projects/knee-segment.

INDEX TERMS Atlas-based segmentation, knee geometry, label fusion, mesh refinement, MRI segmentation, osteoarthritis.

I. INTRODUCTION

The degenerative joint disease, or osteoarthritis, is the most common chronic condition of the joints. It causes pain and mobility limitation leading to reduced independence and degeneration in the quality of life [1]. Most of the existing research has focused on factors associated with the disease, but the limited exploitation of longitudinal data analysis examining the factors associated with disease onset and progression has resulted in poor prevention and treatment interventions. The quantitative assessment of knee pathology caused by injury or disease, such as the estimation of cartilage loss due to osteoarthritis, largely relies on the delineation of the knee compartments, mainly the bones, cartilages and menisci. Magnetic resonance imaging (MRI) can provide the required input since it is characterized by high sensitivity in visualizing soft tissue [2]. However, manual delineation is susceptible to inter- and intra-rater variability. In addition, the three-dimensional (3D) nature of volumetric data hinders the cross-sectional perception in the outlining process, usually performed over sequential 2D scans. In longitudinal studies, automation and reproducibility in segmentation are important factors that reduce the workload of the experts and preserve segmentation consistency, permitting the quantification of subtle changes. The exploitation of image registration techniques, that spatially map the different images in a common (reference) space, can provide information on the localization of the pathology and tissue alteration.

The primary goal of this work concerns the development of an automated method for segmentation of the knee structures from magnetic resonance (MR) images. Many of the proposed MRI-based knee segmentation methods are designed

The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li.

to operate only on specific knee structures, such as the femur and cartilage, with pre-defined characteristics in respect to shape or other descriptors. We build upon a multi-atlas segmentation method developed initially by Wang *et al.* [3] that can generalize on any anatomical structure and that exploits pre-segmented regions of interest to calculate regional likelihoods. The end-goal is to utilize the segmented knee structures for the construction of personalized models than can support finite element analysis (FEA). FEA requires high-quality volumetric meshes with well-shaped elements, such as hexahedral meshes, which are preferred over the tetrahedral meshes for using significantly fewer elements, having a superior convergence rate [4] and less singularities [5]. Although many fast and robust methods exist for the creation of tetrahedral meshes of arbitrary geometries, this is not the case for hexahedral meshes. Hence, we make use of the method presented in [4] to create high quality, smooth hexahedral meshes. The resulting geometries are well-structured, with the cartilages divided into layers so that specific material properties can be potentially attributed to each layer [6].

This work overall aims to establish an automated methodology for creating subject-specific geometries, extracted from medical images, that can be potentially used for modeling, simulation and analysis of the knee mechanics. The main contributions are summarized as follows:

1. A general multi-atlas segmentation scheme utilizing deformable registration and label fusion is combined with an anatomically-adopted mesh refinement algorithm for optimal modelling of the knee from MRI data.

2. The segmentation and mesh generation scheme is not limited to the bones, but is extended to other structures, such as cartilages and menisci with emphasis on improving the geometrical properties required for a subsequent FEA.

3. The effectiveness of the method is illustrated on two different datasets with varying degree of osteoarthritis, in order to investigate whether the segmentation performance is affected by the OA grade.

4. A short taxonomy on representative MRI-based knee segmentation methods is presented to provide some insights on implemented methodologies, amount of data used for validation, targeted structures in the knee complex, obtained accuracy, amount of supervision, and other methodological aspects.

II. REVIEW OF KNEE SEGMENTATION METHODS

A variety of promising methods have been introduced for the segmentation of the knee structures in MR images. The diverse approaches can be categorized coarsely into three groups, namely atlas-based, shape or model-based, and graph-based approaches, without omitting the fact that hybrid methods can also exist. Comparison of the different approaches is not straightforward, since all present advantages and weaknesses in respect to obtained accuracy, requirement of prior information (imaging or segmentation masks), level of automation, sensitivity to initialization, ability to handle abnormal (pathological) cases, computational complexity, etc. Their accuracy also depends strongly on the evaluation data, which might differ in quality and amount, therefore direct comparison is not feasible. A classification scheme of methodologies for the knee joint segmentation was proposed in [7] and also followed in a subsequent review [8], where MRI-based knee bone segmentation methods are analyzed and compared. We extend the previous review by focusing more on the critical structures to be segmented, i.e. cartilages and menisci, and present a table (Table 1) summarizing the main characteristics of current implementations, while in the following we briefly elaborate on methodological aspects.

A main family of methods that has been utilized for knee segmentation relies on statistical shape models (SSMs). The shape is an important characteristic of rigid structures, such as the bones and the cartilages, and can guide the segmentation process by reducing ambiguity. The shapes of the bones can also help identifying weight bearing regions that might cause loss of cartilage in the case of osteoarthritis [9], [10]. The framework in [11] made use of a 3D active shape model for the extraction of the bone-cartilage interface and a subsequent SVM classifier for the extraction of the cartilages. Another strategy utilized active appearance models with the minimum description length criterion for groupwise registration to identify the bones and cartilages [12], while SSMs with multi-object graph optimization were used for the segmentation of bones that assisted the extraction of the cartilages in [13]. Additionally to the segmentation of multiple anatomical compartments in the knee region, shape models were also the dominant choice for the segmentation of specific structures of the knee, such as the menisci [14]–[16]. Furthermore, volumetric appearance models [17] have been used as an extension of surface-based active appearance models for improvement of the segmentation of the knee bones and reduction of the variance in the obtained measurements. Finally, SSMs have been used as an intermediate regularization step between an initial segmentation using 2D convolutional neural networks (CNNs) and an ultimate refinement step with 3D CNNs [16], [18].

Graphical models [19], [20] were also exploited for knee segmentation. They offer a compact image representation and allow to use existing algorithms from the graph theory for image analysis [21]. The segmentation problem is approached by dividing a graph to subgraphs that represent meaningful objects of interest. Specifically, the method in [20] is based on a joint optimization scheme using support vector machines for the description of image features and spatial dependencies and discriminative random fields for the interaction term, while the method in [19] utilizes localized Markov random fields that adaptively emphasize appearance and shape priors, according to local region voxel intensities. Simultaneous segmentation of multiple surfaces, where the segmentation problem is treated as a graph optimization problem, is performed in [22] and recently extended to capture spatiotemporal (longitudinal) context [23]. Content-based features extracted from the edges and the gray level

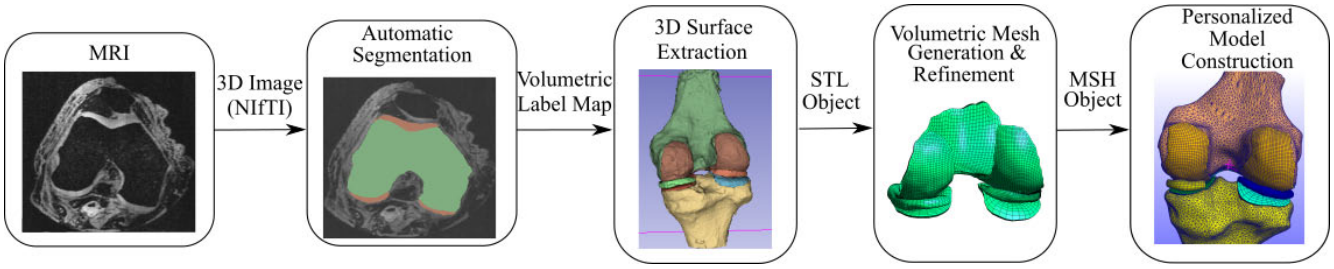


FIGURE 1. Components of our approach towards the generation of a finite element model of the knee complex. After each step the corresponding output file type is given: NIfTI format stands for medical images, STL for 3D surfaces and MSH for volumetric meshes.

co-occurrence matrix have been utilized within a graph-cut segmentation scheme to improve discrimination ability in [24]. Finally, a hybrid multi-atlas and graph-based segmentation approach is adopted in [25].

Although SSMs and graphical models have a strong theoretical basis, they usually require the (joint) optimization of an energy function incorporating an image-based supervised classification cost, therefore increasing computational complexity. Our approach is based on atlas-based segmentation and data fusion. This choice was mainly driven by the high potential of such approaches that has been illustrated in different segmentation problems, along with the recent availability of annotated images of the knee, which is a necessity for atlas-based approaches, and the easier interpretation of the method's intermediate results (in the form of probability maps).

The foundation of atlas-based methods is a registration step, that maps the diverse information from several annotated atlases acquired by a single or multiple imaging techniques into a common space where statistics can be extracted. After spatial normalization each methodology can utilize a different approach to the segmentation process. The initial concept of atlas-based segmentation methodologies relied on single atlases, in which the labels of the atlas were propagated to the novel image via registration. This strategy is highly prone to registration errors and the given atlas might also not be anatomically representative of the novel image. Therefore, the construction of a probabilistic atlas image from a set of atlases started to dominate the field aiming to better capture wide anatomical variations. Multi-atlas segmentation was firstly introduced in a series of papers [26]–[28] that proposed to combine disparate sources of experimental data across multiple targets in order to increase predictive power [29], [30]. Such approaches rely on a fusion scheme, that optimally aggregates the votes of the multiple atlases in order to reach a consensus (final segmentation).

Although multi-atlas segmentation has been applied to a great extend with success for brain image segmentation [31], it has rarely been used for modeling the knee complex. However, multi-atlas approaches have shown great potential in the segmentation of knee bones and cartilages [32], [33], and menisci [34]. The technique proposed in [32] employs outlier detection with voxel classification for the delineation of the cartilages. Class predictions were obtained by each atlas and then combined with a fuzzy voting algorithm. Based on this work, a non-local patch-based label fusion method was realized in [33] coupled with voxel classification to derive the femoral and tibial cartilages. Similarly, profoundly inspired from the work in [35], k-nearest neighbor (kNN) voxel classification was performed after multi-atlas rigid registration to a common space (defined as the training center) in [34]. In the latter work information fusion from the multiple atlases was not performed probabilistically in decision level, but was effectuated implicitly during supervised classification. Such basic data fusion strategies weight equally the contribution of each atlas and moreover do not consider the different levels of uncertainty over the spatial domain. Those important aspects were addressed in [3] and motivated the design of our knee segmentation framework.

Other methods, that rely solely on the intensity content, have also been proposed for MRI-based knee structures segmentation and can involve texture analysis [36], voxel classification with kNN classifiers (without spatial alignment) [35], regression [37] or deep learning [38]–[41]. The deep learning approaches applied 2D CNNs over single [39] or multiple planes [38], or accounted for shape preservation through 3D simplex deformable modeling [40]. The latter work was extended in [41] by incorporating also a 3D fully connected conditional random field in order to consider the 3D contextual relationship among voxels. To avoid segmentation inconsistencies across slices, 3D convolution was performed in [42] using a network inspired by the U-Net architecture [43]. In most of the cases the deep learning approaches have shown superior potential, especially when 3D topology was considered. However, deep networks require the availability of a plethora of segmented images to be used for training purposes. On the opposite, atlas-based segmentation methods are much more robust to a limited number of annotated data.

III. METHODS

The overall scheme includes several methodological components as depicted in Fig. 1. First MRI segmentation is performed based on deformable registration with multiple atlases and label fusion to automatically extract the various knee structures. Then the 3D surface is extracted from the segmentation mask, followed by volumetric mesh generation and refinement of the individual structures (cartilages and menisci). The output consists of a personalized model of the

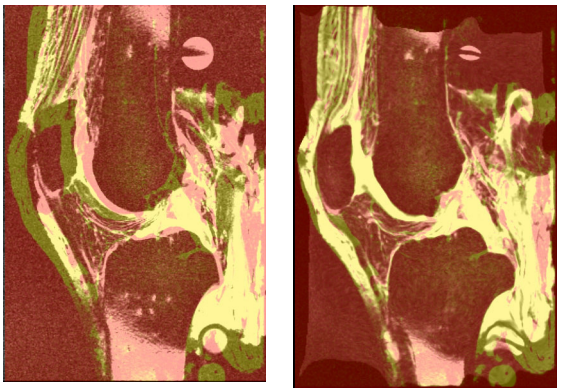


FIGURE 2. Overlay of the target image on the atlas a) before and b) after deformable registration. The soft tissues of the atlas are depicted with a green hue, whereas the soft tissues of the target image with a white hue.

knee complex that is in appropriate form (i.e. is characterized by high mesh quality without degenerated elements) to be used for finite element analysis.

A. REGISTRATION

Atlas-based segmentation methods require that the atlas images and the new unlabeled image are in spatial alignment. This is achieved by deformable registration in which a non-linear transformation is sought that brings each atlas and target image in the same space so that the corresponding anatomical structures spatially coincide. Before performing non-linear transformation, a rigid or affine transformation with a few degrees of freedom is required to correct for global pose and scale differences. In other words, rigid registration achieves the coarse alignment of the images, allowing the subsequent deformable registration to capture the local individual anatomical variations.

Fig. 2 shows an example of a target image overlaid on the atlas before and after deformable registration. The lack of complete overlap of the two structures after registration is mainly explained by the regularization (smoothness) constraints inherent in every deformable registration algorithm. Regularization is necessary because it allows to preserve the physical properties of the underlying anatomical structures by controlling the properties (e.g. compressibility) of the deformation field, and also helps the optimization process to avoid local minima. The residual variability of the coregistered tructures is exploited by the subsequent fusion mechanism, as explained in the next section. Overall, the atlas-based segmentation methodology consists of multiple steps that are illustrated in Fig. 3 and outlined below:

1. The atlas images and the target image are registered in order to achieve anatomical correspondence. This can be achieved in two ways: either by registering all the atlases onto the target image space (let's denote it with Ω_T), or by registering all the atlases and the target image into a common reference atlas (e.g. one of the atlases) and hence work in a *template space* Ω_A. In the latter case a transformation $\Phi : \Omega_T \rightarrow \Omega_A$ is sought and applied to warp the target image into the common space.

2. The warped atlases provide information on the possible localization of the different anatomical structures in the reference space. This information is combined through a label fusion technique to derive the segmentation of the target image and the spatial probability maps for each structure.

3. If the reference space is not the *target image space* the obtained segmentation can be mapped back into the original image space through the application of the inverse transformation Φ^{-1}.

Two methods were mainly exploited for the registration step, namely Elastix [44], which is computationally efficient, and DRAMMS [45], which is more costly but offers more accurate alignment of the images due to its hierarchical attribute matching mechanism and also provides more precise inversion of the deformation fields. Aiming at accuracy, we selected DRAMMS and decided to work in a common reference space in which all atlas images are warped only once, i.e. there is no need to repeat this operation for new-coming images. However, in the case of registration in the target image space, where the multiple atlas images have to be co-registered at testing time, Elastix was the only computationally feasible option. Specifically, the computational cost for Elastix was 30s and for DRAMMS 108 minutes in an Intel Core i7-4770 CPU @ 3.40 GHz x 8 with a 16 GB RAM machine. Since DRAMMS performs only deformable registration, the initial global alignment of the images was performed using FLIRT [46] of the FSL[1] library. The FLIRT tool was used with a 6 degrees of freedom (DOF) transformation model (i.e. rigid registration) and the normalized correlation as similarity metric. In a few cases where linear registration was poor, we augmented the transformation model to 12 DOF (i.e. affine registration).

B. MULTI-ATLAS SEGMENTATION BY LABEL FUSION

Segmentation of the multiple knee structures was performed by adopting the technique described in [3] that optimizes the fusion of segmentation outcomes from multiple atlases through a mechanism referred by the authors as *joint label fusion*. In this fusion mechanism, the contribution of each atlas is weighted according to the probability that each pair of atlases makes a segmentation error at the specific location and then weighted voting is applied to obtain the final consensus vote. The weights do not only depend on each atlas' uncertainty, but also follow a spatially varying distribution based on the intensity similarity between atlas and target image.

With more details, let T_I be the target intensity image, T_S the corresponding unknown segmentation map and $A^1 = \{A_I^1, A_S^1\}, \ldots, A^n = \{A_I^n, A_S^n\}$ the n atlases (each including a pair of intensity and ground truth segmentation images), warped to the space of the T_I image with deformable registration. We suppose that we have D image modalities and L

[1] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSL

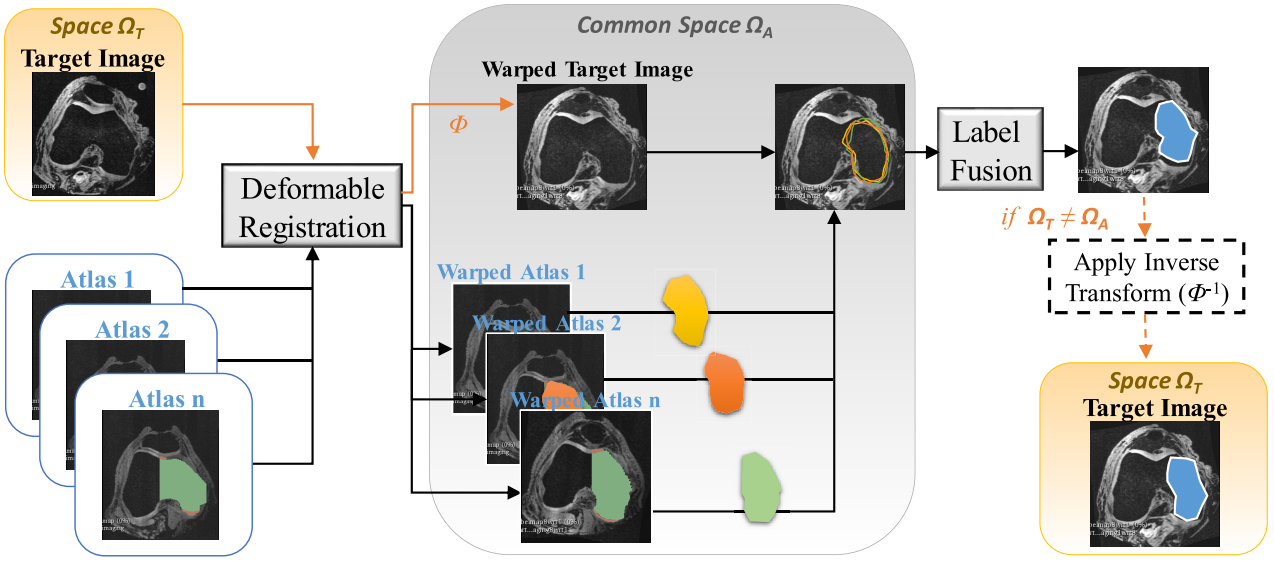


FIGURE 3. The automatic segmentation scheme. Atlases and target image are warped in a common space Ω_A through deformable registration to achieve anatomical correspondence. If the "template space" where label fusion is performed, is not the original space of the target image (Ω_T), the output segmentation is mapped back to the target image space.

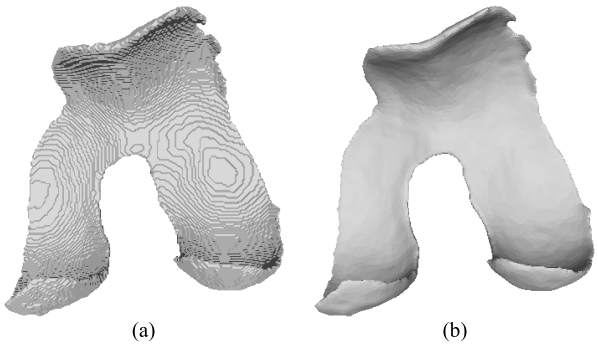


FIGURE 4. a) 3D surface of the femoral cartilage from the marching cubes algorithm; b) The same surface after the application of a Laplacian filter.

different labels (regions of interest). For deterministic atlases, where each image position has a unique label (hard segmentation), the estimated probability to receive a consensus vote for label l at location x of the target image T_I is:

$$\hat{p}\left(l \mid x, T_I\right) = \sum_{i=1}^{n} w_x^i I\left(A_S^i\left(x\right) = l\right) \tag{1}$$

where $I\left(A_S^i\left(x\right) = l\right)$ is an indicator function with the value of 1 if the atlas A^i was segmented with label l at location x and 0 otherwise, and w_x^i is the voting weight for the atlas A^i with $\sum_{i=1}^{n} w_x^i = 1$.

In order to estimate the weights w_x^i, the method does not treat each atlas independently from the others but takes into account that different atlases might produce similar segmentation errors. The segmentation errors produced by each registered atlas are modeled by $\delta^i\left(x\right) \in \{-1, 0, 1\}$, with 0 indicating label agreement, i.e., that both (T_S and A_S^i) or neither one has the label l at voxel x, while -1 or 1 indicates label mismatch, i.e., that only one of the two segmentations has label l at voxel x. The probability that two atlases produce the same label error at position x is modeled by the dependency matrix $M_x\left(i,j\right) = p\left(\delta^i\left(x\right)\delta^j\left(x\right) = 1 \mid T_I, A_I^i, A_I^j\right)$. Since the segmentation of the target image is unknown, to assess the segmentation outcome, the similarity of the intensity images is used as an indicator of the segmentation correctness. This is based on the principle that when two patches in the intensity images have similar profiles, they most probably should have similar segmentation. To minimize the expected label difference between the consensus solution obtained from weighted voting and the target segmentation, the optimal voting weights are determined by

$$\boldsymbol{w}_x = \frac{\boldsymbol{M}_x^{-1}\mathbf{1}_{\boldsymbol{n}}}{\mathbf{1}_{\boldsymbol{n}}^t\boldsymbol{M}_x^{-1}\mathbf{1}_{\boldsymbol{n}}} \tag{2}$$

where $\mathbf{1}_{\boldsymbol{n}} = [1, 1, \ldots, 1]^t$ is a vector of size n. To avoid inverting an ill-conditioned matrix an identity matrix weighted by a small positive number α is added to $\boldsymbol{M}_x$. The dependency matrix is estimated over local image patches and depends on the intensity difference between the target image and each atlas aggregated over all image modalities.

Similarly, the label fusion is not performed over the whole image context, but over local image patches. The pair of image patches is determined after a local search of correspondence. Specifically, based on the premise that when two image regions look similar, they are more likely to correspond to the same structure (segmentation label), if $N\left(x\right)$ is the patch to be segmented in the target image, the method searches in each atlas i for the most similar location $\xi_i\left(x\right)$, within a neighborhood N_s:

$$\xi_i\left(x\right) = \arg\min_{x' \in N_s(x)} \left\| A_I^i\left(N\left(x'\right)\right) - T_I\left(N\left(x\right)\right)\right\|^2. \tag{3}$$

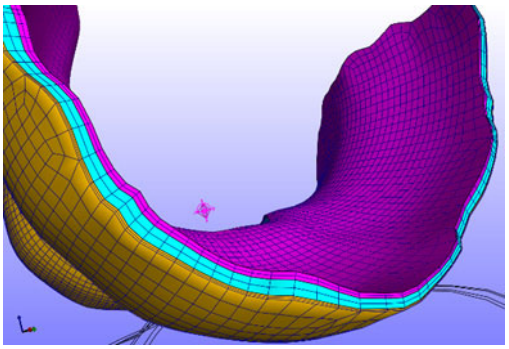


FIGURE 5. Femoral cartilage layers grouped into three zones that permits the assignment of different material properties to model different cellular synthesis.

This local search makes the method more robust to registration errors. Accordingly, the consensus vote received by label l in location x is updated to include the (spatially varying) weights at the location $\xi_i(x)$ of the corresponding atlas patch

$$\hat{p}(l \mid x, T_I) = \sum_{i=1}^{n} w^i(\xi_i(x)) I\left(A_S^i(\xi_i(x)) = l\right). \quad (4)$$

The above method also incorporates a machine learning strategy to detect and correct the systematic errors produced by any incorporated segmentation method. This latter strategy, which is called *corrective learning* in [3], is based on the hypothesis that a large part of the segmentation errors produced by an automatic method occur consistently from subject to subject. This step in the methodology attempts to learn the intensity, spatial and contextual patterns associated with systematic segmentation errors produced by the host method on training data (the atlases) for which manual segmentation is available. The method then attempts to correct errors in the segmentation produced by the host method on new images.

C. MORPHOLOGICAL PROCESSING

The output of the multi-atlas segmentation technique might contain errors that result in loss of the shape regularity. A frequent case is a segmentation with isolated pieces. To resolve this issue, the isolated segments were removed by keeping the largest connected component of each structure using 6-connectivity as the elementary 3D neighborhood of the morphological operation [47]. To avoid the removal of segments that lie close to the largest connected component and probably should be part of the final structure, a morphological image closing operation is first performed for each label.

IV. MESH GENERATION AND REFINEMENT

After performing the segmentation, the surface mesh can be extracted from the 3D discrete scalar field using a surface polygonization technique, such as the very popular *marching cubes* algorithm [48]. Since the marching cubes algorithm produces a rather rough surface a Laplacian filter was applied to smooth the geometries (Fig. 4) [49]. To reduce

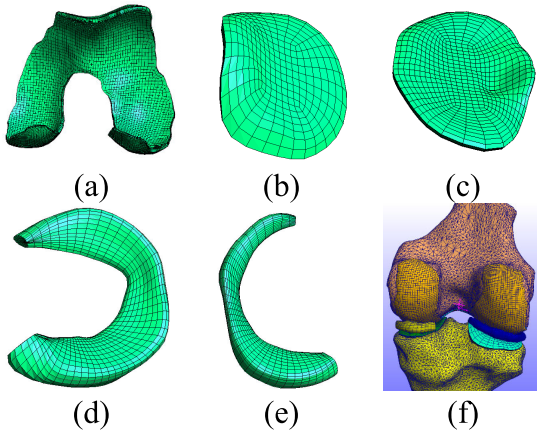


FIGURE 6. Results of the mesh generation algorithm: a) femoral cartilage, b) lateral tibial cartilage, c) medial tibial cartilage, d) lateral meniscus, e) medial meniscus and f) assembled finite element model.

the computational cost of the subsequent hexahedral meshing algorithm, the faces of the 3D surfaces were limited to 10000 using the Quadric Edge Collapse Decimation filter [49].

Next, the previous triangular meshes were converted to hexahedral meshes [4]. For this purpose, an anatomically-adopted sweeping algorithm is applied to obtain a low-resolution mesh with hexahedral elements followed by Laplacian smoothing to improve the mesh quality after the segmentation. As a quality measure of the hexahedral mesh, the scaled Jacobian metric is used. Specifically, the nodal position n_j of each node j of the low-resolution mesh is updated based on local information using an extended version of the Laplacian smoothing algorithm [50] as

$$\widehat{\boldsymbol{n}}_j = \alpha \boldsymbol{n}_j + \frac{1-\alpha}{\|adj(j)\|} \sum_{k \in adj(j)} \boldsymbol{n}_k \quad (5)$$

where $adj(j)$ are the adjacent nodes of node j, and α is a parameter controlling the balance between smoothing quality and degree of model shrinkage ($\alpha = 0.5$ was used). Since the Laplacian filter shrinks the input geometry's shape, the smoothed mesh is expanded towards the triangular surface. The mesh is optimized by moving the nodes from their respective positions in order to obtain a proper Jacobian for the elements [51]. Smoothing, expansion and optimization are iteratively repeated until the quality metric is met [52] and the elements are ultimately subdivided to formulate 6 layers for the cartilages (Fig. 5) or a more refined mesh for the menisci [4]. The output of the hexahedral meshing algorithm is shown in Fig. 6.

V. RESULTS

The method was evaluated on two distinct data sets. The first dataset included MR images from 8 subjects as part of the Open Knee project [53]. One subject was excluded since the MRI modalities deviated significantly from the other samples, therefore 7 subjects we included in our analysis.

The data were used in a cross-validation setting for training and testing. The MRI protocol is referred as ''Cartilage Imaging'', which is a T1-weighted with fat saturation sequence, providing good contrast for the cartilages. The Open Knee project supplements high quality ground-truth masks that were used to guide the segmentation method. These masks do not only include the bones and the cartilages, but also the menisci (along with other ligaments) that potentially offer the chance for future extension of this work to other structures as well. The second dataset included MR images (3D DESS WE) from an initial set of 78 subjects randomly selected from the Osteoarthritis Initiative (OAI) public use data set. The OAI dataset does not provide ground truth masks for the menisci that are required for the reconstruction process; therefore, this dataset was not used in the assessment of the mesh refinement technique.

The MR images were segmented to obtain subject-specific geometries that are required for the construction of personalized models. For the Open Knee dataset multi-atlas registration was performed in the target image space [44] due to the small number of subjects, whereas for the OAI dataset the reference space was retained fixed to a common atlas space [45]. Three metrics were employed to evaluate the accuracy of the segmentation, namely the Dice similarity coefficient (DSC), the false negative ratio (FNR) and the mean surface distance (MSD). The latter is also referred in the literature as average distance (AvgD). The performance of the hexahedral meshing algorithm was also investigated to assess whether it modifies significantly the output volumetric geometries with respect to the input geometries. The 3D meshes were extracted and evaluated with respect to the input surfaces, using the MSD metric.

The abbreviations in the illustrated results (Fig. 7 to Fig. 9) are as follows: FB – femur bone, TB – tibia bone, FC – femoral cartilage, LTC – lateral tibial cartilage, MTC – medial tibial cartilage, TC – tibial cartilage (lateral & medial as one structure), LM – lateral meniscus and MM – medial meniscus. Finally, we investigated whether our results could be exploited for extraction of imaging biomarkers for OA assessment. For this purpose, we investigated if there is any correlation between the OA grade and the medial cartilage volume of the tibia, as will be explained later.

A. ASSESSMENT OF THE SEGMENTATION AND MESH GENERATION USING THE OPEN KNEE DATASET

This section presents the results of the segmentation accuracy on the Open Knee dataset, using three quality metrics (DSC, FNR and MSD), in order to facilitate comparison with the segmentation approaches of other authors (Table 1). Two segmentation approaches were examined, one including the final machine learning scheme (corrective learning) and one by replacing this computationally expensive scheme with simple morphological operations. The results were cross-validated with the leave-one-out strategy, i.e., the images of 6 subjects were used as atlases and the left-out subject as target image.

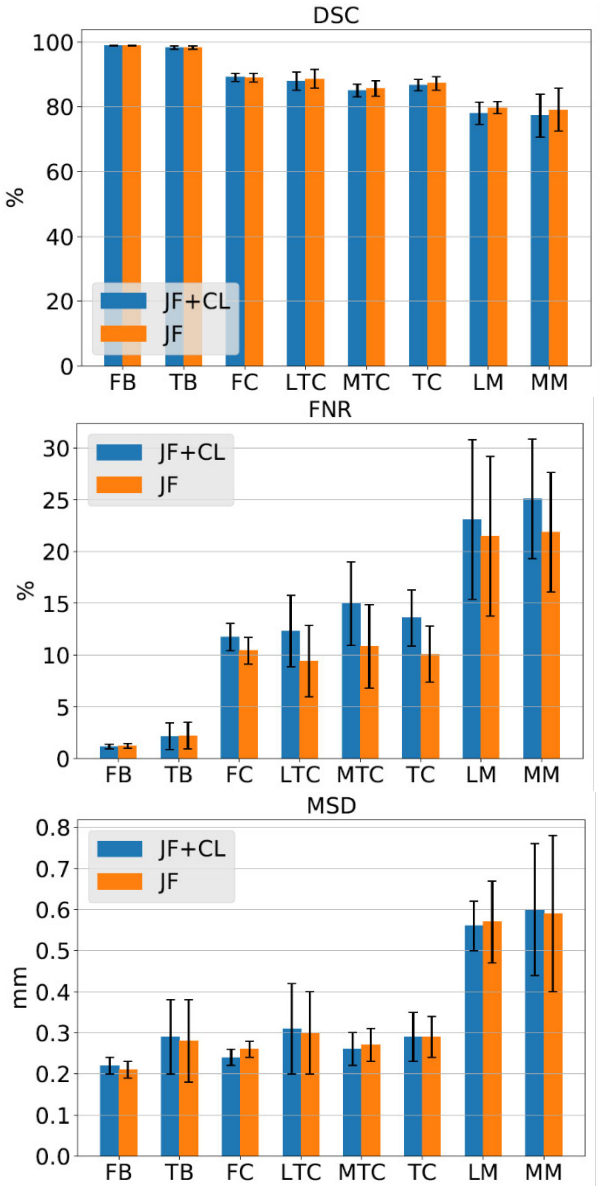


FIGURE 7. Bar plots of the DSC, FNR and MSD similarity metrics for the segmentation of the Open Knee data, for two case scenarios, namely (i) using the corrective learning scheme (JF+CL) and (ii) based solely on JF. The JF approach achieves slightly higher accuracy as displayed in the DSC graph and more clearly in the FNR graph, but no conclusion can be derived from the MSD graph. FB – femur bone, TB – tibia bone, FC – femoral cartilage, LTC – lateral tibial cartilage, MTC – medial tibial cartilage, TC – tibial cartilage (lateral & medial as one structure), LM – lateral meniscus and MM – medial meniscus.

1) ASSESSMENT OF CORRECTIVE LEARNING

Fig. 7 depicts several similarity scores for two approaches, namely one with and one without the corrective learning (CL) technique [3] that is designed to detect systematic segmentation errors. By comparing the performance of these two approaches, we observe that the segmentation (in Fig. 7) obtained solely by multi-atlas joint label fusion (JF) displays slightly higher DSC and lower FNR. By inspecting the DSC scores the maximum difference appears in the case

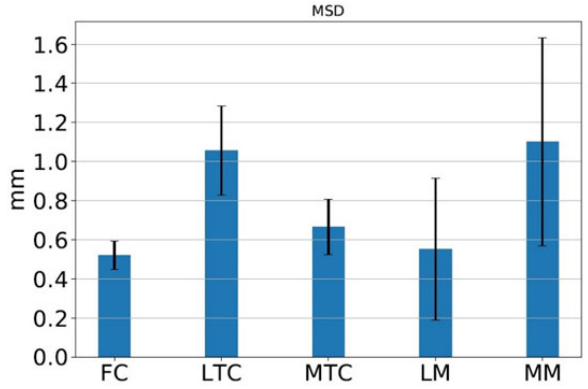


FIGURE 8. Mean and standard deviation of the MSD (mm) between the contour of volumetric mesh and the corresponding initial 3D surface (Open Knee data). The largest difference occurs for the medial meniscus (MM) and has a mean value of 1.1 mm.

of lateral meniscus where the segmentation accuracy is 2.14 % higher than the one obtained by CL. Larger differences are observed using the FNR metric (lower FNR signifies better performance). The maximum difference appears for the medial tibial cartilage (27.7% higher FNR with CL) with the remaining cartilages and menisci showing improvement by omitting the corrective learning step, as well. On the contrary, not all the MSD scores display a benefit loss from the CL approach. The maximum absolute difference appears for the femoral cartilage (8.3% higher accuracy). Hence, no conclusion can be drawn based on the MSD score. The small number of available cases in the Open Knee project might not be sufficient to postulate that the incorporation of the CL step has any benefit on the accuracy. Therefore, since the marginal difference in accuracy does not justify the significant increase in computational cost by the CL step, for the proposed modeling scheme we consider only the joint label fusion as default step for image segmentation.

2) ASSESSMENT OF MORPHOLOGICAL PROCESSING

We separately evaluated the effect of the morphological operations performed after the joint label fusion technique (without CL) in order to preserve the shape topology. The DSC has improved for all structures after morphological processing, although the differences were quite small. Specifically, for the femur and tibia bones the differences in DSC were insignificant (<0.1), while for FC, LTC, MTC, LM and MM the DSC has increased by 0.23, 0.50, 0.16, 0.35 and 0.20, respectively. In respect to FNR, as expected the removal of isolated pieces and the retainment of the largest connected component could only cause an increase in FNR, since this metric takes into account only the false negatives. Therefore, this metric was not considered the most representative for assessment. In any case, the increase in FNR values was too small (<0.7) to dominate over the necessity to enforce shape regularity and was outreached by the improvement of DSC.

3) COMPARISON WITH OTHERS

In Table 1 a comparison of different segmentation methods is presented. Although direct comparison is not feasible due

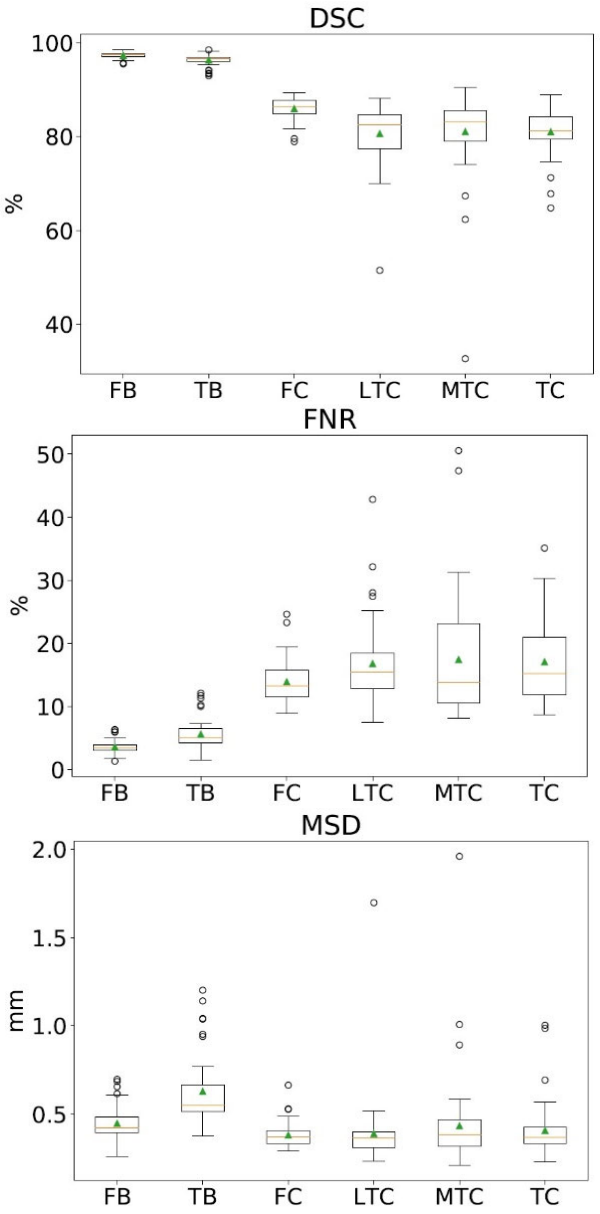


FIGURE 9. Box plots displaying the segmentation accuracy of the OAI data. For the cases with extensive outliers it would be preferable to consider the median value in place of the mean value.

to differences in the datasets, we overall observe that the implemented JF segmentation method outperforms most of the other methods, in terms of higher DSC/MSD scores. Additionally, this method achieved comparable accuracy in some cases, e.g. in terms of DSC score, 1.4% difference for the tibia bone in [33], 0.8% for the tibial cartilage in [20] and 5% for the lateral meniscus in [15], [34]. The method employed in [18] achieved comparable segmentation accuracy for the cartilages in terms of the DSC score, but this is not the case for the MSD scores, where [18] gives 38% and 20.7% more accurate results for the femoral and tibial cartilages. When the previous method is adopted for the menisci [16], it gives 12% higher DSC score for the lateral meniscus and

6% for the medial meniscus. The results of Fig. 7 are also presented with numeric values in Table 2 in the Appendix.

In addition to the previous comparative analysis, we evaluated the performance of an open source knee segmentation workflow, the pyKNEEr [54]. For fair comparison, we selected the OpenKnee dataset due to its high quality ground-truth masks and adjusted all the images to have the same right laterality by flipping the left knee images into the right knee images, as we had performed also for the evaluation of our method. In pyKNEEr each new image is registered into the atlas space using Elastix [44] and the reference mask is warped onto the target image through the inverse transformation field. The segmentation step is based on a single atlas and focuses only on the segmentation of the femoral cartilage. In fact, the femur bone is initially segmented in order to guide the segmentation of the femoral cartilage. The obtained DSC, calculated through leave-one-out cross validation, was very low, i.e. 61.2% for the femur bone and 49% for the femoral cartilage. These results show the benefit of the multi-atlas label fusion approach incorporated in our segmentation scheme.

4) ASSESSMENT OF MESH GENERATION

The generation of volumetric meshes was performed only for the Open Knee data since this dataset includes ground truth segmentations of the menisci. The mean and standard deviation values between the initial 3D surface geometries and the contour of the resulting volumetric geometries are provided in Fig. 8. Out of all structures, the surface difference is larger for the medial meniscus, but still relatively small ($1.1 \pm 0.53\ mm$).

B. SEGMENTATION ASSESSMENT ON THE OAI DATASET

In this section, the proposed segmentation method was applied to a larger dataset of knee MR images with varying degrees of OA and differences in the images' intensity distributions, to test whether these factors affect the accuracy of the method. Eighty subjects were initially randomly selected, but two subjects were excluded from the analysis because the Kellgren and Lawrence score, assessing the OA degree (required in our next experiment), could not be retrieved for the one and the other resulted in especially high cartilage volume ($> 3000\ mm^3$) in the common atlas space and was considered a rare case (outlier). The selection is randomized with 33 subjects being used for training and 45 for testing. Segmentation was applied in a common template space (all the images were registered to a reference atlas), in which results were evaluated without being mapped back in the original domain. The default JF approach was used for the segmentation without the incorporation of the CL method.

By comparing the DSC, FNR and MSD scores with the ones obtained from the Open Knee data it is obvious that the accuracy is slightly smaller, e.g. for the femoral and tibial cartilage the DSC scores are 86.01% and 81.04% for the OAI, with 88.96% and 87.26% for the Open Knee respectively.

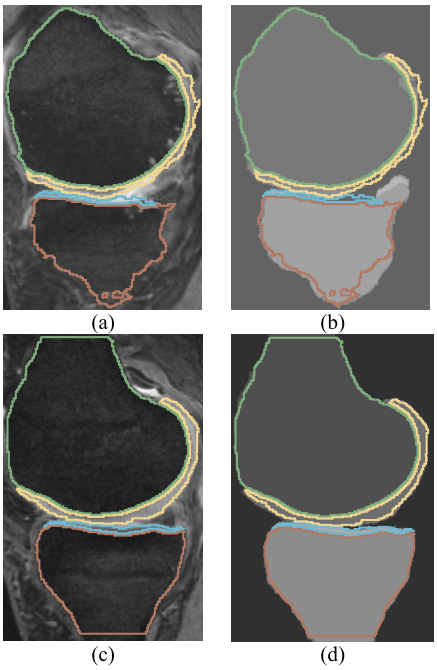


FIGURE 10. Example of a bad (1st row) and good (2nd row) automatic segmentation result. a-b) Bad segmentation overlaid on the MRI (left) and the ground truth segmentation (right), c-d) good segmentation overlaid on the MRI (left), and the ground truth segmentation (right).

By examining the femoral and tibial cartilages, the proposed segmentation yields scores that are superior over several works e.g. [12], [19], [22], [25], [35] (Table 1). The works [20], [32] achieved high scores for the tibial cartilage (8.6% and 3.7% superior), as well for the femoral cartilage in [32] (2.2% superior), but using a method [20] that relies on multiple modalities, that are not always available. Specifically, for the tibial cartilage, we achieved 81.04% mean DSC score, which is comparable to the works [32], [33], [36], [38] (maximum DSC 84.1%) and within the range of scores obtained by [11], [34]. The performance of these two methods [11], [34] was low for the femoral cartilage (DSC $< 84.4\%$), except for the method by Tamez-Pena *et al.* [32] where the accuracy for the femoral cartilage was comparable with the proposed method (88% DSC versus 86.01% for our segmentation). Since the data employed in this work are a subset of the OAI ZIB dataset, we observe that the recent work by Ambellan *et al.* [18] (which used the whole dataset) results in higher DSC scores (by 4.5% for the femoral cartilage and 5.6% for the tibial cartilage) than the proposed method. Fig. 9 displays the box plots for the 45 testing subjects using the same similarity metrics as for the Open Knee data.

As observed, there exist several cases (outliers) with large deviations from the mean values of DSC and FNR for the tibial cartilages (lateral and medial). The mean values of Fig. 9 are provided in Table 3 with supplementary similarity metrics (Fig. 12). Fig. 10 illustrates a case with bad and a case with good DSC scores. In the first case we observe that low image registration quality, possibly due to large anatomical

TABLE 1. Summary of state-of-the-art methods used for MRI-based knee segmentation.

	Atlas based	Training data	No. of subjects	Structure	Accuracy	Supervised	Shape model based	Graph based
Proposed	✓	✓	85 (Open Knee, OAI)	FB TB FC LTC MTC TC LM^2 MM^2	DSC*: 97.41 - 98.84 DSC: 96.42 - 98.24 DSC: 86.01 - 88.96 DSC: 80.66 - 88.60 DSC: 81.08 - 85.64 DSC: 81.04 - 87.26 DSC: 79.69 DSC: 79.11	✓		
Lee et al. 2014 [25]	✓	✓	150 SKI10	FB TB FC TC	AvgD**: 0.63 RMSD**: 1.05 AvgD: 0.53 RMSD: 0.9 DSC: 71.7 VOE*:28.3 VD*: -5.5 DSC: 72.4 VOE: 27.6 VD:-3.7	✓		✓
Tamez – Pena et al. 2012 [32]	✓	✓	48 MRI	FC TC	DSC: 88.0 DSC: 84.0	✓		
Shan et al. 2014 [33]	✓	✓	706 PLS for 115 subjects & 50 SKI10	FB TB FC TC	DSC: 96.9 DSC: 99.6 DSC: 76.0 DSC: 84.1	✓		
Dam et al. 2015 [34]	✓	✓	1907 (CCBR, OAI, VirtualScopics, Chondometrics)	FC TC LM MM	DSC: 80.4 – 84.4 DSC: 80.5 – 86.6 DSC: 82.2 - 83.0 DSC: 76.0	✓		
Dodin et al. 2010 [36]			14 MRI (9 standard, 5 test/retest)	FC TC	DSC: 84.0 DSC: 82.0			
Fripp et al. 2010 [11]		✓	20 MRI	PC FC TC	DSC: 73.0 - 83.0 DSC: 79.0 - 83.0 DSC: 76.0 - 86.0	✓		
Lee et al. 2011 [19]		✓	17 MRI	FB TB PB FC TC PC	DSC: 95.2 DSC: 96.4 DSC: 95.4 DSC: 82.0 DSC: 81.0 DSC: 82.0	✓	✓	✓
Seim et al. 2010 [13]		✓	40 MRI	FB TB FC TC	AvgD: 1.02 RMSD: 1.54 AvgD: 0.84 RMSD: 1.24 VOE: 34.0 VD: 7.7 VOE: 29.2 VD: -2.7	✓	✓	✓
Vincent et al. 2010 [12]			80 OAI	FB TB FC	AvgD: 0.88 RMSD: 1.49 AvgD: 0.74 RMSD: 1.21 VOE: 36.3	✓	✓	

TABLE 1. ***(Continued.)*** **Summary of state-of-the-art methods used for MRI-based knee segmentation.**

			TC	VD: -25.2 VOE: 34.6 VD: -9.5			
Zhang et al. 2013 [20]	✓	11 MRI	FC TC PC	DSC: 86.4 DSC: 88.0 DSC: 84.1	✓		✓
Prasoon et al. 2013 [38]	✓	114 MRI	TC	DSC: 82.49	✓		
Yin et al. 2010 [22]	✓	60 OAI	FC TC PC	DSC: 84.0 DSC: 80.0 DSC: 80.0	✓	✓	✓
Williams et al. 2010 [17]	✓	48 MRI from 12 subjects	restricted to central, load-bearing regions	not clearly defined	✓	✓	
Folkesson et al. 2007 [35]	✓	139 MRI	FC TC	DSC: 77.0 DSC: 81.0	✓		
Ababneh et al. 2011 [24]	✓	376 MRI	Mean bone	DSC: 95.0	✓		✓
Ambellan et al. 2019 [18]	✓	88 OAI Imorphics	FC LTC MTC	DSC: 89.4 MSD: 0.19 DSC: 90.4 MSD: 0.17 DSC: 86.1 MSD: 0.26	✓	✓	
		507 OAI ZIB	FB TB FC TC	DSC: 98.6 MSD: 0.17 DSC: 98.5 MSD: 0.18 DSC: 89.9 MSD: 0.16 DSC: 85.6 MSD: 0.23			
		150 SKI10	FB TB FC TC	RMSD: 0.74 RMSD: 0.59 VD: 7.18 VOE: 20.99 VD: 4.29 VOE: 19.06			
Tack et al. 2018 [16]	✓	88 OAI Imorphics	LM MM	DSC: 88.9 DSC: 83.8	✓	✓	
Paproki et al. 2014 [15]	✓	88 OAI Imorphics	LM MM	DSC: 83.9 DSC: 78.3	✓	✓	
Fripp et al. 2019 [14]	✓	14 MRI	LM MM	DSC: 77.0 DSC: 75.0	✓	✓	
Norman et al. 2018 [39]	✓	174 OAI (70/20/10 train/validation/ test)	FC LTC MTC PC LM MM	DSC: 86.7 DSC: 79.9 DSC: 77.7 DSC: 76.7 DSC: 81.2 DSC: 73.1	✓		
Liu et al. 2018 [40]	✓	100 SKI10 (60/40 train/test)	FB TB FC TC	AvgD: 0.56 RMSD: 1.08 AvgD: 0.50 RMSD: 1.09 VD: 8.1 VOE: 28.4 VD: -1.2 VOE: 33.1	✓		

TABLE 1. ***(Continued.)*** **Summary of state-of-the-art methods used for MRI-based knee segmentation.**

Zhou et al. 2018 [41]	✓	20 MRI	FB TB PB FC TC PC LM+MM	DSC: 97.0 DSC: 96.2 DSC: 89.8 DSC: 80.6 DSC: 80.1 DSC: 80.7 DSC: 83.1	✓	✓
Raj et al. 2018 [42]	✓	100 SKI10 (80/20 train/test)	FC TC	DSC: 83.4 VOE: 28.3 DSC: 82.5 VOE: 29.5	✓	

The studies are classified in respect to the implemented methodology, amount of data, targeted structures, obtained accuracy, amount of supervision, etc. VOE – volumetric overlap error, VD – volumetric difference, RMSD – root mean square symmetric surface distance; FB – femur bone, TB – tibia bone, PB – patella bone, FC – femoral cartilage, LTC – lateral tibial cartilage, MTC – medial tibial cartilage, TC – tibial cartilage (lateral & medial as one structure), PC – patellar cartilage, LM – lateral meniscus, MM – medial meniscus. * measured in percentage (%). ** measured in millimeter (mm).2 Lateral and medial meniscus segmentation was available only from the Open Knee project.

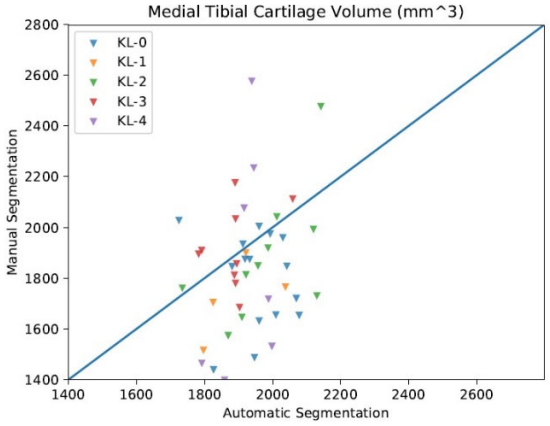


FIGURE 11. **Scatter plots of Kellgren and Lawrence (KL) score (0 healthy – 4 severe OA) for the medial tibial cartilage extracted from manual and automatic segmentation.**

differences, resulted in poor segmentation of the test image. However, even though the DSC score is small, visual inspection reveals a reasonable segmentation result that indicates robust outcomes. The good segmentation case (on the 2nd row) displays much smoother segmentation boundaries that accurately circumscribe the various structures.

C. EXPLOITATION OF OUR RESULTS FOR MORPHOMETRIC QUANTIFICATION

Next, we wanted to investigate whether our results could be used to extract imaging features for OA assessment. The OA grade was assessed by the Kellgren and Lawrence (KL) score, ranging from 0 for a healthy knee to 4 for severe OA. Generally, the entire context of the knee joint should be examined to assess the OA degree. However, the clinical symptoms are dominating in the medial compartments of tibial plateau [55], [56]. Thus, for the analysis, we only examined the volume of the medial tibial cartilage, similarly to previous work [57].

In order to examine possible sources of error introduced by our automatic approach, we first compared the automatically extracted MTC volumes with the ones based on the expert segmentation. The results are plotted in Fig. 11 and allow to assess whether there is any systematic error in the segmentation method with respect to the cartilage volume or the osteoarthritis grade. We don't observe any evident trend in the distribution of errors in volume estimation (automatic vs manual) and the medial tibial cartilage volume itself. Moreover, the difference in volume estimation between the two methods does not seem to be directly related to the OA grade. However, the volumes in cases with lower OA grades tend to be overestimated by the automatic method (i.e. the majority of cases with KL 0 lie below the main diagonal in the graph). Further investigation of whether image artifacts (caused by disease deterioration) affect the segmentation accuracy could not be performed due to the inability to retrieve follow up studies, that included manually annotated data and OA grade, for the OAI dataset.

In respect to morphometric quantification, neither segmentation method revealed a significant correlation of MTC volume with KL score at the baseline examination ($R = 0.14$, p-value $= 0.35$ for the automatic method and $R = 0.01$, p-value $= 0.93$ for the manual segmentation). This outcome, which contradicts previous studies on morphometric cartilage quantification [57], might be due to the fact that we did not consider differences on total knee size as confounding factor. Quantitative work in the future will include a standardization step, in which for example the measurements will be normalized in respect to the width of the tibial plateau, in order to account for knee size differences related to age or sex.

Moreover, we will include shape-related measures, such as cartilage thickness and curvature, which proved to be more sensitive for separation of the healthy from low grade OA [57], [58].

VI. DISCUSSION

Studies have shown that MRI has good diagnostic performance in detecting OA [59] due to the high soft tissue

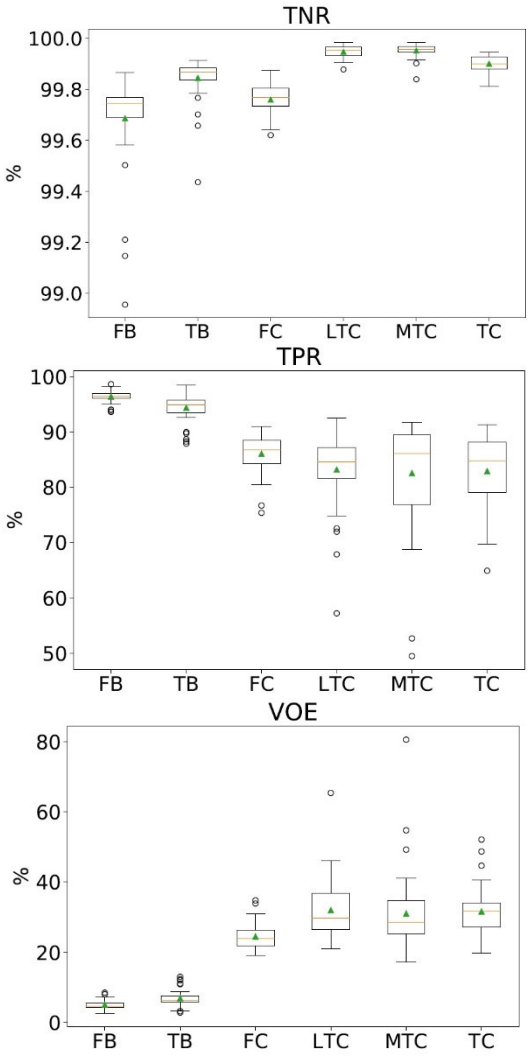


FIGURE 12. Box plots of the TNR, TPR and VOE similarity metrics for the OAI data to allow for comparison with works employing metrics other than DSC, FNR, MSD.

TABLE 2. Performance metrics[a] (mean and standard deviation) for the Open Knee data using the multi-atlas segmentation method (Joint Fusion – JF) with and without the corrective learning (CL) technique.

	DSC (%)		FNR (%)		MSD (mm)	
	JF+CL	JF	JF+CL	JF	JF+CL	JF
FB	98.81 ± 0.11	98.84 ± 0.12	1.15 ± 0.25	1.23 ± 0.23	0.22 ± 0.02	0.21 ± 0.02
TB	98.21 ± 0.50	98.24 ± 0.57	2.13 ± 1.08	2.20 ± 1.29	0.29 ± 0.09	0.28 ± 0.10
FC	89.05 ± 1.33	88.96 ± 1.34	11.73 ± 1.19	10.43 ± 1.30	0.24 ± 0.02	0.26 ± 0.02
LTC	87.89 ± 2.75	88.60 ± 2.83	12.34 ± 4.63	9.42 ± 3.45	0.31 ± 0.11	0.30 ± 0.10
MTC	85.01 ± 1.92	85.64 ± 2.32	14.97 ± 6.86	10.83 ± 4.01	0.26 ± 0.04	0.27 ± 0.04
TC	86.62 ± 1.73	87.26 ± 2.05	13.58 ± 3.58	10.08 ± 2.7	0.29 ± 0.06	0.29 ± 0.05
LM	78.02 ± 3.46	79.69 ± 1.82	23.10 ± 8.16	21.48 ± 7.71	0.56 ± 0.06	0.57 ± 0.10
MM	77.31 ± 6.60	79.11 ± 6.61	25.10 ± 5.68	21.85 ± 5.76	0.60 ± 0.16	0.59 ± 0.19

[a]DSC (Dice Similarity Coefficient), FNR (False Negative Rate), MSD (Mean Surface Distance, same as Average Surface Distance - AvgD)

contrast and the ability to track histological changes, such as bone marrow lesions, joint fluid changes, and cartilage loss [60]. A mechanism modeling the tissue alteration caused by the disease could prove to be beneficial in understanding the pathophysiological process of OA histological changes. We propose a modular MRI-based methodology that allows the automatic quantification of the tissue in the multiple knee compartments facilitating reproducible analysis and consistent follow-up studies.

The proposed method demonstrated high segmentation accuracy for the case of Open Knee data as compared to most of the state-of-the-art techniques (summarized in Table 1), whereas for the OAI data the segmentation performance was comparable to the other methods. This might be attributed to the high quality of the Open Knee segmentation masks that offer more accurate delineation of the underlying structures as well as the common imaging properties (intensity distribution) of the utilized MR images, which were all acquired with the same protocol. On the other hand, the limited number of subjects in the Open Knee dataset does not allow us to generalize our conclusions. Overall, exact comparison with the works of the others is not feasible due to differences in the incorporated datasets. However, by examining the literature on methods using small datasets [14], [19], [20], [36] it appears that the proposed segmentation scheme achieves better performance, except of the methodology presented in [16], [18] which seems to outperform our method.

The main reason for exploiting the Open Knee dataset was the availability of segmented menisci along with the bones and cartilages. It is essential to note that the employed imaging protocol does not offer good contrast for the menisci (or other soft tissues) as would be the case for a proton density (PD) MR protocol. Since the incorporated segmentation method is modular to the imaging modalities, if PD images are available in the future, they can be easily combined to obtain a more accurate segmentation of the menisci.

As for the performance of the hexahedral meshing algorithm, we observe (Fig. 8) that the volumetric geometries do not deviate substantially from the original 3D surfaces (scores derived from cross correlation). The largest MSD value tends to occur on the medial meniscus, but the small size of the dataset does not provide sufficient evidence to assume that the MSD score of this structure is consistently higher. An advantage of the anatomically-adopted algorithms for the generation of the hexahedral meshes is that the structure is divided into layers. This feature can be useful to model different cellular synthesis, e.g. progression of diseases such as the OA, by assigning different material properties to each layer [61].

VII. CONCLUSION

The aim of this work is to provide an end-to-end pipeline for automatic segmentation and 3D reconstruction of the

TABLE 3. Performance metrics[a] (mean and standard deviation) for the OAI dataset.

	DSC (%)	FNR (%)	TNR (%)
FB	97.41 ± 0.65	3.62 ± 1.07	99.69 ± 0.17
TB	96.42 ± 1.19	5.64 ± 2.39	99.85 ± 0.08
FC	86.01 ± 2.41	13.95 ± 3.41	99.76 ± 0.06
LTC	80.66 ± 6.42	16.82 ± 6.56	99.95 ± 0.02
MTC	81.08 ± 9.12	17.47 ± 9.51	99.95 ± 0.02
TC	81.04 ± 4.79	17.10 ± 6.42	99.9 ± 0.03
	TPR (%)	VOE (%)	MSD (mm)
FB	96.38 ± 1.07	5.04 ± 1.23	0.447 ± 0.099
TB	94.36 ± 2.39	6.89 ± 2.2	0.628 ± 0.186
FC	86.05 ± 3.41	24.46 ± 3.65	0.381 ± 0.071
LTC	83.18 ± 6.55	31.97 ± 8.23	0.389 ± 0.208
MTC	82.53 ± 9.51	31.01 ± 10.58	0.433 ± 0.273
TC	82.90 ± 6.42	31.61 ± 6.48	0.407 ± 0.150

[a]DSC (Dice Similarity Coefficient), FNR (False Negative Rate), TNR (True Negative Rate), TPR (True Positive Rate), VOE (Volumetric Overlap Error), MSD (Mean Surface Distance, same as Average Surface Distance - AvgD)

multiple knee compartments from MRI. The geometrical structures can be subsequently introduced into biomechanical simulation models for patient-specific estimation of articular loading, or for the calculation of the external knee adduction moment required during gait retraining interventions. The presented scheme is modular with respect to the exploited imaging modalities. It achieved competitive segmentation performance and produced high quality meshes. Such computer modeling approaches, once combined with simulations aggregating various information sets from full body, organ, or tissue level mechanistic models, open the path for the prediction of optimal treatments, better diagnostics, and improved outcomes of patients with degenerative joint disease.

APPENDIX

See Figure 12 and Tables 2 and 3.

ACKNOWLEDGMENT

The authors would like to thank K. Filip for his help with data preparation and constructive discussions on the various topics of this work. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

[1] A. Mündermann, C. O. Dyrby, and T. P. Andriacchi, "A comparison of measuring mechanical axis alignment using three-dimensional position capture with skin markers and radiographic measurements in patients with bilateral medial compartment knee osteoarthritis," *Knee*, vol. 15, no. 6, pp. 480–485, Dec. 2008.

[2] F. Eckstein, F. Cicuttini, J.-P. Raynauld, J. C. Waterton, and C. Peterfy, "Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): Morphological assessment," *Osteoarthritis Cartilage*, vol. 14, pp. 46–75, Jan. 2006.

[3] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 611–623, Mar. 2013.

[4] B. Rodriguez-Vila, P. Sánchez-González, I. Oropesa, E. J. Gomez, and D. M. Pierce, "Automated hexahedral meshing of knee cartilage structures—Application to data from the osteoarthritis initiative," *Comput. Methods Biomech. Biomed. Eng.*, vol. 20, no. 14, pp. 1543–1553, Oct. 2017.

[5] S. C. Tadepalli, A. Erdemir, and P. R. Cavanagh, "Comparison of hexahedral and tetrahedral elements in finite element analysis of the foot and footwear," *J. Biomech*, vol. 44, no. 12, pp. 2337–2343, Aug. 2011.

[6] K. S. Halonen, C. M. Dzialo, M. Mannisi, M. S. Venäläinen, M. de Zee, and M. S. Andersen, "Workflow assessing the effect of gait alterations on stresses in the medial tibial cartilage–combined musculoskeletal modelling and finite element analysis," *Sci. Rep.*, vol. 7, no. 1, pp. 1–14, Dec. 2017.

[7] Y. Sun, E. C. Teo, and Q. H. Zhang, "Discussions of Knee joint segmentation," in *Proc. Int. Conf. Biomed. Pharmaceutical Eng.*, Dec. 2006.

[8] A. Aprovitola and L. Gallo, "Knee bone segmentation from MRI: A classification and literature review," *Biocybernetics Biomed. Eng.*, vol. 36, no. 2, pp. 437–449, 2016.

[9] J. G. Tamez-Peña, M. Barbu-McInnis, and S. Totterman, "Unsupervised definition of the tibia-femoral joint regions of the human knee and its applications to cartilage analysis," in *Proc. Med. Imag., Image Process.*, vol. 6144, Mar. 2006, p. 61444.

[10] T. G. Williams, "Cartilage loss in osteoarthritis detected by statistical shape analysis of magnetic resonance images," *Osteoarthr. Cartil.*, vol. 13, p. 228, 2005.

[11] J. Fripp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 55–64, Jan. 2010.

[12] G. Vincent, C. Wolstenholme, I. Scott, and M. Bowes, "Fully automatic segmentation of the knee joint using active appearance models," *Med. Image Anal. Clin. Grand Challenge*, pp. 224–230, Sep. 2010.

[13] H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, "Model-based auto-segmentation of knee bones and cartilage in MRI data," *Med. Image Anal. Clin. Grand Challenge*, pp. 215–223, Nov. 2010.

[14] J. Fripp, P. Bourgeat, C. Engstrom, S. Ourselin, S. Crozier, and O. Salvado, "Automated segmentation of the menisci from MR images," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano Macro*, Jun. 2009, pp. 510–513.

[15] A. Paproki, C. Engstrom, S. S. Chandra, A. Neubert, J. Fripp, and S. Crozier, "Automated segmentation and analysis of normal and osteoarthritic knee menisci from magnetic resonance images–data from the osteoarthritis initiative," *Osteoarthritis Cartilage*, vol. 22, no. 9, pp. 1259–1270, Sep. 2014.

[16] A. Tack, A. Mukhopadhyay, and S. Zachow, "Knee menisci segmentation using convolutional neural networks: Data from the osteoarthritis initiative," *Osteoarthritis Cartilage*, vol. 26, no. 5, pp. 680–688, May 2018.

[17] T. G. Williams, "Automatic segmentation of bones and inter-image anatomical correspondence by volumetric statistical modelling of knee MRI," in *Proc. 7th IEEE Int. Symp. Biomed. Imag. Nano Macro (ISBI)*, Apr. 2010, pp. 432–435.

[18] F. Ambellan, A. Tack, M. Ehlke, and S. Zachow, "Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: Data from the osteoarthritis initiative," *Med. Image Anal.*, vol. 52, pp. 109–118, Feb. 2019.

[19] S. Lee, S. H. Park, H. Shim, I. D. Yun, and S. U. Lee, "Optimization of local shape and appearance probabilities for segmentation of knee cartilage in 3-D MR images," *Comput. Vis. Image Understand.*, vol. 115, no. 12, pp. 1710–1720, Dec. 2011.

[20] K. Zhang, W. Lu, and P. Marziliano, "Automatic knee cartilage segmentation from multi-contrast MR images using support vector machine classification with spatial dependencies," *Magn. Reson. Imag.*, vol. 31, no. 10, pp. 1731–1743, Dec. 2013.

[21] N. Paragios, E. Ferrante, B. Glocker, N. Komodakis, S. Parisot, and E. I. Zacharaki, "(Hyper)-graphical models in biomedical image analysis," *Med. Image Anal.*, vol. 33, pp. 102–106, Oct. 2016.

[22] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka, "LOGISMOS—Layered optimal graph image segmentation of multiple objects and surfaces: Cartilage segmentation in the knee joint," *IEEE Trans. Med. Imag.*, vol. 29, no. 12, pp. 2023–2037, Dec. 2010.

[23] S. Kashyap, H. Zhang, K. Rao, and M. Sonka, "Learning-based cost functions for 3-D and 4-D multi-surface multi-object segmentation of knee MRI: Data from the osteoarthritis initiative," *IEEE Trans. Med. Imag.*, vol. 37, no. 5, pp. 1103–1113, May 2018.

[24] S. Y. Ababneh, J. W. Prescott, and M. N. Gurcan, "Automatic graph-cut based segmentation of bones from knee magnetic resonance images for osteoarthritis research," *Med. Image Anal.*, vol. 15, no. 4, pp. 438–448, Aug. 2011.

[25] J.-G. Lee, S. Gumus, C. H. Moon, C. K. Kwoh, and K. T. Bae, "Fully automated segmentation of cartilage from the MR images of knee using a multi-atlas and local structural analysis method," *Med. Phys.*, vol. 41, no. 9, Aug. 2014, Art. no. 092303.

[26] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NeuroImage*, vol. 33, no. 1, pp. 115–126, Oct. 2006.

[27] A. Klein, B. Mensh, S. Ghosh, J. Tourville, and J. Hirsch, "Mindboggle: Automated brain labeling with multiple atlases," *BMC Med. Imag.*, vol. 5, no. 1, pp. 1–14, Dec. 2005.

[28] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *NeuroImage*, vol. 21, no. 4, pp. 1428–1442, Apr. 2004.

[29] S. Alchatzidis, A. Sotiras, E. I. Zacharaki, and N. Paragios, "A discrete MRF framework for integrated multi-atlas registration and segmentation," *Int. J. Comput. Vis.*, vol. 121, no. 1, pp. 169–181, Jan. 2017.

[30] M. B. Vakalopoulou, G. Chassagnon, N. Bus, and R. Marini, in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, vol. 8150. Springer, 2013.

[31] J. E. Iglesias and M. R. Sabuncu, "Multi-atlas segmentation of biomedical images: A survey," *Med. Image Anal.*, vol. 24, no. 1, pp. 205–219, Aug. 2015.

[32] J. G. Tamez-Pena, J. Farber, P. C. Gonzalez, E. Schreyer, E. Schneider, and S. Totterman, "Unsupervised segmentation and quantification of anatomical knee features: Data from the osteoarthritis initiative," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp. 1177–1186, Apr. 2012.

[33] L. Shan, C. Zach, C. Charles, and M. Niethammer, "Automatic atlas-based three-label cartilage segmentation from MR knee images," *Med. Image Anal.*, vol. 18, no. 7, pp. 1233–1246, Oct. 2014.

[34] E. B. Dam, M. Lillholm, J. Marques, and M. Nielsen, "Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative," *J. Med. Imag.*, vol. 2, no. 2, Apr. 2015, Art. no. 024001.

[35] J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen, "Segmenting articular cartilage automatically using a voxel classification approach," *IEEE Trans. Med. Imag.*, vol. 26, no. 1, pp. 106–115, Jan. 2007.

[36] P. Dodin, J. Pelletier, J. Martel-Pelletier, and F. Abram, "Automatic human knee cartilage segmentation from 3-D magnetic resonance images," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 11, pp. 2699–2711, Nov. 2010.

[37] A. Saygılı and S. Albayrak, "A new computer-based approach for fully automated segmentation of knee meniscus from magnetic resonance images," *Biocybernetics Biomed. Eng.*, vol. 37, no. 3, pp. 432–442, 2017.

[38] A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a Triplanar convolutional neural network," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, no. 9. 2013, pp. 246–253.

[39] B. Norman, V. Pedoia, and S. Majumdar, "Use of 2D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry," *Radiology*, vol. 288, no. 1, pp. 177–185, Jul. 2018.

[40] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao, and R. Kijowski, "Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging," *Magn. Reson. Med.*, vol. 79, no. 4, pp. 2379–2391, Apr. 2018.

[41] Z. Zhou, G. Zhao, R. Kijowski, and F. Liu, "Deep convolutional neural network for segmentation of knee joint anatomy," *Magn. Reson. Med.*, vol. 80, no. 6, pp. 2759–2770, Dec. 2018.

[42] A. Raj, S. Vishwanathan, B. Ajani, K. Krishnan, and H. Agarwal, "Automatic knee cartilage segmentation using fully volumetric convolutional neural networks for evaluation of osteoarthritis," in *Proc. Int. Symp. Biomed. Imag.*, Apr. 2018, pp. 851–854.

[43] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.

[44] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.

[45] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," *Med. Image Anal.*, vol. 15, no. 4, pp. 622–639, 2011.

[46] B. Fischer and J. Modersitzki, "FLIRT: A flexible image registration toolbox," in *Proc. Int. Workshop Biomed. Image Registration*, 2010, pp. 261–270.

[47] V. G. Kanas, E. I. Zacharaki, C. Davatzikos, K. N. Sgarbas, and V. Megalooikonomou, "A low cost approach for brain tumor segmentation based on intensity modeling and 3D random walker," *Biomed. Signal Process. Control*, vol. 22, pp. 19–30, Sep. 2015.

[48] B. R. de Araújo, D. S. Lopes, P. Jepp, J. A. Jorge, and B. Wyvill, "A survey on implicit surface polygonization," *ACM Comput. Surveys*, vol. 47, no. 4, pp. 1–39, May 2015.

[49] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "Meshlab: An open-source mesh processing tool," in *Proc. Eurographics Italian Chapter Conf.*, 2008, pp. 129–136.

[50] J. Vollmer, R. Mencl, and H. Müller, "Improved Laplacian smoothing of noisy surface meshes," *Comput. Graph. Forum*, vol. 18, no. 3, pp. 131–138, Sep. 1999.

[51] M. Auer and T. C. Gasser, "Reconstruction and finite element mesh generation of abdominal aortic aneurysms from computerized tomography angiography data with minimal user interactions," *IEEE Trans. Med. Imag.*, vol. 29, no. 4, pp. 1022–1028, Apr. 2010.

[52] W. B. Lievers and R. W. Kent, "Patient-specific modelling of the foot: Automated hexahedral meshing of the bones," *Comput. Methods Biomech. Biomed. Eng.*, vol. 16, no. 12, pp. 1287–1297, Dec. 2013.

[53] A. Erdemir and S. Sibole, "Open knee: A three-dimensional finite element representation of the knee joint," vol. 25, Dec. 2015.

[54] S. Bonaretti, G. E. Gold, and G. S. Beaupre, "PyKNEEr: An image analysis workflow for open and reproducible research on femoral knee cartilage," *PLoS ONE*, vol. 15, no. 1, 2019, Art. no. e0226501.

[55] T. C. Dunn, Y. Lu, H. Jin, M. D. Ries, and S. Majumdar, "T2 relaxation time of cartilage at MR imaging: Comparison with severity of knee osteoarthritis," *Radiology*, vol. 232, no. 2, pp. 592–598, 2004.

[56] L. Kamibayashi, U. P. Wyss, T. D. V. Cooke, and B. Zee, "Changes in mean trabecular orientation in the medial condyle of the proximal tibia in osteoarthritis," *Calcified Tissue Int.*, vol. 57, no. 1, pp. 69–73, Jul. 1995.

[57] E. B. Dam, J. Folkesson, P. C. Pettersen, and C. Christiansen, "Automatic morphometric cartilage quantification in the medial tibial plateau from MRI for osteoarthritis grading," *Osteoarthritis Cartilage*, vol. 15, no. 7, pp. 808–818, Jul. 2007.

[58] T. G. Williams, A. P. Holmes, J. C. Waterton, R. A. Maciewicz, C. E. Hutchinson, R. J. Moots, A. F. P. Nash, and C. J. Taylor, "Anatomically corresponded regional analysis of cartilage in asymptomatic and osteoarthritic knees by statistical shape modelling of the bone," *IEEE Trans. Med. Imag.*, vol. 29, no. 8, pp. 1541–1559, Aug. 2010.

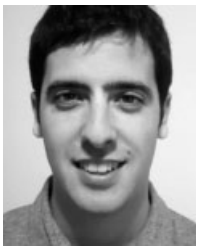
[59] L. Menashe, K. Hirko, E. Losina, M. Kloppenburg, W. Zhang, L. Li, and D. J. Hunter, "The diagnostic performance of MRI in osteoarthritis: A systematic review and meta-analysis," *Osteoarthritis Cartilage*, vol. 20, no. 1, pp. 13–21, Jan. 2012.

[60] L. M. Shapiro, E. J. McWalter, M.-S. Son, M. Levenston, B. A. Hargreaves, and G. E. Gold, "Mechanisms of osteoarthritis in the knee: MR imaging appearance," *J. Magn. Reson. Imag.*, vol. 39, no. 6, pp. 1346–1356, Jun. 2014.

[61] K. Huch, "Knee and ankle: Human joints with different susceptibility to osteoarthritis reveal different cartilage cellularity and matrix synthesis *in vitro*," *Arch. Orthopaedic Trauma Surgery*, vol. 121, no. 6, pp. 301–306, Jun. 2001.

FILIPPOS P. NIKOLOPOULOS received the Diploma degree in electrical and computer engineering and the M.Sc. degree in biomedical engineering from the University of Patras, Greece, in 2017 and 2019, respectively. From 2018 to 2019, he was with the Virtual Reality Group, Electrical and Computer Engineering Department, University of Patras, for the European project OActive that concerned the development of patient-specific computer models for the treatment of Osteoarthritis. His work was focused on the automatic atlas-based segmentation of knee MRI and the modeling and refinement of the 3D mesh of the structures of the knee complex.

EVANGELIA I. ZACHARAKI received the M.Eng., Ph.D., and H.D.R. degrees in 1999, 2004, and 2017, respectively. She worked as a Research Associate with the Section of Biomedical Image Analysis, UPenn, USA, from 2005 to 2009, the Medical Physics Department, University of Patras, from 2009 to 2012, and the Center for Visual Computing, CentraleSupélec/INRIA, France, from 2015 to 2017. She is currently a Senior Research Scientist with the Visualization and Virtual Reality Group, Electrical and Computer Engineering Department, University of Patras, Greece. She has coauthored in total 100 articles in refereed 38 international journals, two books , and 60 conference proceedings and has received around 2060 citations until March 2020 (H-index=20, G-index=43). Her research interests focus on machine learning, computational and statistical modeling for representation, and fusion and analysis of high-dimensional biomedical data. Dr. Zacharaki received a Marie Curie Reintegration Grant and has participated in more than 15 international (NIH/NIA), European (FP7, H2020), and national research projects.



DIMITAR STANEV received the B.S. and Ph.D. degrees in electrical and computer engineering from the University of Patras, Greece, in 2014 and 2018, respectively. He is currently a Postdoctoral Researcher with Biorobotics Laboratory (BioRob), Institute of Bioengineering, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. His research focuses mostly on aspects related to musculoskeletal modeling and simulation, muscle redundancy, proprioception and locomotion, knee mechanics and finite element. Since 2014, he has been participated in several research and development projects funded by the European Commission and the Swiss Government. He is a proponent of the open science movement with a contribution to the OpenSim project, the development of open access models and simulation methods. Since 2019, he has been named an inaugural OpenSim Fellow of the NIH National Center for Simulation in Rehabilitation Research (NCSRR).

KONSTANTINOS MOUSTAKAS (Senior Member, IEEE) received the Diploma and Ph.D. degrees in electrical and computer engineering from the Aristotle University of Thessaloniki, Greece, in 2003 and 2007, respectively. From 2007 to 2011, he served as a Postdoctoral Research Fellow with the Information Technologies Institute, Centre for Research and Technology Hellas. He is currently an Associate Professor with the Electrical and Computer Engineering Department, University of Patras, and the Head of the Visualization and Virtual Reality Group. He is also the Director of the Wire Communications Laboratory and the Master's Program Biomedical Engineering of the University of Patras. During the latest years, he has been the (co)author of more than 180 articles in refereed journals, edited books, and international conferences. He serves as a regular reviewer for several technical journals and has participated to more than 20 research and development projects funded by the EC and the Greek Secretariat of Research and Technology. He was the coordinator of the GameCar Horizon2020 project, scientific coordinator of the NoTremor FP7 project, while he also chaired the scientific board of the myAirCoach and FrailSafe H2020 projects. His main research interests include virtual, augmented and mixed reality, 3D geometry processing, virtual physiological human modeling, biomedical engineering, information visualization, and physics-based simulations.