Deep Learning-Based Segmentation of Mineralized Cartilage and Bone in High-Resolution Micro-CT Images

Jean Léger[1(✉)], Lisa Leyssens[2], Christophe De Vleeschouwer[1], and Greet Kerckhofs[2,3,4,5]

[1] ICTEAM, UCLouvain, Louvain-la-Neuve, Belgium
jean.leger@uclouvain.be
[2] iMMC, UCLouvain, Louvain-la-Neuve, Belgium
[3] Institute of Experimental and Clinical Research, UCLouvain, Louvain-la-Neuve, Belgium
[4] Department of Materials Science and Engineering, KU Leuven, Leuven, Belgium
[5] Prometheus, Division of Skeletal Tissue Engineering, KU Leuven, Leuven, Belgium

Abstract. High-resolution 3D micro-CT imaging is a powerful tool for the visualization of the mineralized tissues. However, it remains challenging to discriminate automatically between mineralized cartilage and bone as they have similar greyscale values. Currently, manual contouring is still the standard way to segment these two tissues but it is time-consuming and user-biased. In this work, we have optimized a 3D fully convolutional neural network, i.e. U-net, to automatically segment mineralized cartilage from bone in high-resolution micro-CT images of the Achilles tendon-to-bone interface. Using the 3D U-net, we reach an average Dice Similarity Coefficient of 0.85 compared to manual annotations for twelve 3D datasets. The proposed method shows comparable results to a 2D U-net approach while ensuring better 3D segmentation consistency. We also found that reducing the resolution of the 3D micro-CT images for the network training did not importantly impact the performance while considerably reducing the training time.

Keywords: Mineralized tissue segmentation · CNN · Deep learning · Cartilage versus bone · High-resolution micro-CT

1 Introduction

The musculoskeletal system supports and stabilizes the human body and coordinates the movements of the muscles and the skeletal system. It is composed of bone, muscle, cartilage, tendon, ligament, and other connective tissues. In particular, tendons and bone are joined in a specific way in order to facilitate joint motion, forming the insertion site [1]. This is also known as the enthesis or the bone-to-tendon interface. The tissue that makes up this interface is a complex transitional tissue, which is essential for physiologic musculoskeletal motion.

G. A. Ateshian et al. (Eds.): CMBBE 2019, LNCVB 36, pp. 158–170, 2020.
https://doi.org/10.1007/978-3-030-43195-2_12

It serves to integrate and minimize stress concentrations between the bone (a stiff, structural, hard tissue) and the tendon (a compliant structural soft tissue). More precisely, the bone-to-tendon interface can be divided into four specific zones with varying compositional and mechanical properties and functions [2]: the tendon, the non-mineralized fibrocartilage, the mineralized fibrocartilage, and the bone.

Because of its function of mediating load between very dissimilar tissues, the bone-to-tendon interface is a common site of injury. However, in case of such injury, the natural tissue is not regenerated after healing [1,3]. Instead, it is replaced with a so-called scar tissue, which is isotropic and significantly less stiff than either tendon or bone. It has been reported that depending on the severity and the location of the injury, the regenerated tissue will rupture again in 20 to 94% of cases [4]. A solution to this problem would be to use a construct composed of a scaffold (made of a biomaterial) with growth factor and/or cells to obtain a tissue that presents the same properties as the original one.

In this context, characterizing the bone-to-tendon interface properties and its 3D sub-architecture is primordial. Towards this goal, high-resolution 3D contrast-enhanced microfocus computed tomography (CE-CT) is, thanks to its high spatial and contrast resolution, a powerful tool for the visualization of both the unmineralized and mineralized cartilage, along with the bone [5]. However, it remains challenging to discriminate automatically between the mineralized cartilage and the bone in both high-resolution micro-CT and CE-CT images, as they have similar greyscale values. Indeed, when using only greyscale-based segmentation, it is very difficult to accurately detect the boundaries between these two mineralized tissues. Currently, manual delineation of the two tissues is still used to allow for its 3D structural analysis and that of the bone. Nevertheless, this is time consuming and highly user-biased.

Interestingly, mineralized cartilage has a discriminant porous texture in comparison to bone, because of the presence of chondrocyte lacunae. Consequently, our goal is to develop an automatic mineralized cartilage segmentation tool exploiting its texture, shape and 3D consistency. For this purpose, we use a fully convolutional neural network [6], named U-net, which is state-of-the-art for the semantic segmentation in 2D and 3D biomedical images [7,8]. The 3D U-net model is compared to the 2D U-net model and the impact of reducing the 3D image resolution for GPU memory limitations is discussed.

To the best of our knowledge, no fully automatic segmentation algorithms have been proposed for the segmentation of mineralized cartilage from bone. Indeed, all the related works mentioned below focus on the easier unmineralized cartilage segmentation. Most works focus on the unmineralized cartilage versus bone segmentation in MR images. To this end, classical medical image segmentation approaches relying on active shape models [9] and atlas databases [10] have been proposed. However, those methods perform poorly in case of high ROIs shape variability and require a relatively long segmentation times at inference. In contrast, deep learning algorithms are supposed to be robust to shape and appearance variations if those variations are captured in the train-

ing database and they are fast at inference. In a first study on deep-learning based knee cartilage segmentation, they used a triplanar CNN [11]. This was followed by 2D encoder-decoder architectures such as Segnet [12], U-net [13] and U-net complemented with conditional random fields to promote 3D consistency [14]. Very recent papers use 3D approaches for knee cartilage segmentation, such as 3D U-net [15] or a variant of it, namely μ-Net [16]. Only few methods have been proposed to automatically segment unmineralized cartilage on CT images. They include an atlas-based segmentation of acetabular cartilage [17] and a registration-based segmentation of knee articular cartilage [18] on CE-CT images. One CNN-based unmineralized cartilage segmentation on high-resolution micro-CT images has been proposed. It uses a 2D U-net model followed by a 3D refinement. However, the method has only been presented in a one-page abstract and is therefore hard to reproduce [19].

2 Materials and Methods

2.1 Data Acquisition, Annotation and Preprocessing

The training and test data for the U-net consist of 3D high-resolution micro-CT images of 12 murine bone-to-Achilles tendon interfaces. After harvest, all the samples were fixed in 4% paraformaldehyde during 16 h and then stored at 4 °C in a phosphate-buffered saline solution (PBS). Next, the samples were dissected in order to isolate the tendon and the bone, and to remove all tissue that was unnecessary for the purpose of the experiments. A Phoenix Nanotom M - Computed Tomography System (GE Sensing & Inspection Technologies GmbH, Wunstorf, Germany) was used with the following scanning parameters: 60 kV, 87 µA, 1.25 µm voxel size, 2400 images, 500 ms exposure time, 20 min scan time.

The mineralized cartilage has been manually annotated on all the slices of the micro-CT images by an expert using the CTAn software from Bruker MicroCT (Kontich, Belgium). Those annotations serve as ground truth in this work and they are stored as 12 3D binary masks with the same size and resolution as the 3D micro-CT images.

The original isotropic voxel size is 1.25 µm. The 12 3D datasets have been cropped to 1024 × 1024 × 256 voxels around the region of interest (i.e. mineralized cartilage and bone). The ROI is fully contained in every 1024 × 1024 slice. The bottom slices contain only mineralized cartilage (and no bone) whereas mineralized cartilage and bone start to be difficult to distinguish visually on the upper slices. This preprocessing is necessary since the memory limitations of the GPUs is a bottleneck for deep learning algorithms with such high-resolution data. Nvidia 1080Ti 11 GB GPUs are used.

2.2 Network Architecture

The 3D U-net fully convolutional neural network considered in this study is the same as in Brion et al. [8]. More precisely, the network follows the same

architecture (i.e. number and composition of layers) as in Ronneberger et al. [20], where the 3×3 convolutions, the 2×2 max-pooling and the 2×2 up-conversion operations have been replaced by their $3 \times 3 \times 3$ and $2 \times 2 \times 2$ counterparts, as in Çiçek et al. [6]. The 3D input goes through a contracting path to capture context and an expanding path to enable precise localization. In the last layer, a sigmoid is applied and the network outputs the probability for each voxel to belong to the mineralized cartilage. To obtain the final binary segmentation mask, a threshold of 0.5 is chosen. The main advantage of fully convolutional neural networks is that they output a prediction with the same size as the input. Hence, the thresholded output of the network directly gives the desired prediction of the ROI with the same size and resolution as the input 3D data.

The network is trained with the Dice loss. The optimization algorithm used is Adam with learning rate 10^{-4}. The number of epochs is chosen such that convergence is reached and without using early-stopping. The hyper-parameters mentioned here are the same as in [8] and showed to be satisfactory on the data used in this work. For this reason, no validation set is considered here. We only used two sets: a training set and a test set both containing 6 3D images. The batch size depends on the learning strategy that is used, but has not been tuned.

The results of 3D U-net are compared to those of 2D U-net. The 2D architecture is exactly the same as the 3D one presented above, where the $3 \times 3 \times 3$ convolutions, the $2 \times 2 \times 2$ max-pooling and the $2 \times 2 \times 2$ up-conversion operations have been replaced by their 3×3 and 2×2 counterparts. Online data augmentation (i.e. rotation, shift, shear, flip) is implemented for both 2D and 3D approaches.

2.3 Learning Strategies

Two major interests of the considered datasets are (i) their high-resolution allowing an accurate tissue visualization, in particular for the porous texture characterizing the mineralized cartilage region, and (ii) their ability to capture the 3D architecture of the imaged tissues in a uniformly sampled 3D grid, in opposition to histology where only stacks of 2D slices are available. However, such 3D large high-resolution data are particularly challenging for deep learning models in terms of training time and GPU memory limitations. In order to deal with those constraints, three strategies could be applied.

1. The high-resolution 3D image is uniformly downsampled in every dimension. A 3D model can then be trained and tested on the full 3D image directly. This solves the GPU memory limitation at the expense of a lower resolution.
2. The 3D image is split in 2D slices along a chosen dimension. In this case, a 2D model can be trained and tested at high-resolution, but the 3D consistency is no more ensured.
3. The high-resolution 3D image is split in 3D patches without reducing their resolution. However, this approach requires many training patches in order to capture the variability of the full 3D image. This increases considerably the training time.

In this work, the 12 1024 × 1024 × 256 voxels 3D datasets have been downsampled by a factor 4 in every dimension, resulting in 256 × 256 × 64 voxels 3D datasets. The downsampling factor has been chosen in order to visually preserve the porous texture of the mineralized cartilage. The same operations are performed on the annotation 3D binary masks. A 3D U-net model is then trained on 6 of those low resolution 3D datasets and tested on the 6 remaining ones. In order to determine whether the 3D approach improves the segmentation performances compared to the 2D approach, a 2D U-net model has been trained on 384 low resolution images (6 3D datasets multiplied by 64 slices in each 3D dataset) of 256 × 256 pixels. The 2D test predictions are stacked in order to reconstruct a 3D binary mask and compare it with the 3D binary mask predicted with the 3D approach. In order to evaluate the impact of reducing the spatial image resolution, a 3D U-net model has also been trained on 96 high-resolution patches (16 patches on each of the 6 3D datasets) of 512 × 512 × 16 voxels. The sampling of the patches in the 1024 × 1024 × 256 high-resolution 3D datasets is the following. The patches are sampled uniformly in the third 3D image dimension (i.e. a patch every 16 slices in order to cover the 256 slices) and randomly in the two first dimensions, provided that at least one voxel belongs to the cartilage in the patch. The batch size is set to one for the 3D approaches and 64 for the 2D approach. This is the maximum batch size allowed by the GPU memory limitation. A summary of the three strategies properties is provided in Table 1.

Table 1. Characteristics and training hyper-parameters for the three learning strategies.

Parameters	Strategy 1	Strategy 2	Strategy 3
Model	3D U-net	2D U-net	3D U-net
Input data size	256 × 256 × 64	256 × 256	512 × 512 × 16
Batch size	1	64	1
Number of training samples	6	384	96
Training time per epoch	∼40 s	∼3 s	∼640 s
Number of epochs	∼150	∼150	∼100

2.4 Performance Assessment

We use a two-fold cross-validation (with 6 3D datasets in each fold) in order to obtain the predictions on the 12 3D datasets. The absence of a validation set is motivated in Sect. 2.2. In order to evaluate our results, we use the Dice similarity coefficient (DSC), which measures the overlap between two binary masks. More specifically,

$$\mathrm{DSC} = \frac{2|A \cap B|}{|A| + |B|}, \tag{1}$$

where A and B are the predicted and ground truth segmentation binary masks. The DSC reaches respectively zero and one for no or complete overlap between both binary masks. The low resolution ROIs predictions of the deep learning model data are upsampled by a factor 4 before the DSC computation. Hence, the DSC are computed in the original high-resolution.

3 Results and Discussion

3.1 Quantitative Comparison of the Learning Strategies

Table 2 presents the segmentation performances obtained with the three considered learning strategies. The 3D U-net approach trained on the low resolution datasets reaches an average DSC close to 0.85. This corresponds to visually acceptable segmentation outputs as shown in Figs. 1 and 2, which respectively present the results on Set 2 (DSC = 0.889) and Set 3 (DSC = 0.836). The red and blue regions correspond to the mineralized cartilage ROIs. In particular, in Fig. 1d, the 3D U-net model is able to capture the complex mineralized cartilage structure, while reasonably avoiding the bone. Three types of errors are mostly present on the predicted ROIs.

1. *Errors on the mineralized cartilage extremities.* This can be observed in Figs. 1e and 2e. The extremities of the ROIs are too long compared to the manual annotations. A noteworthy point is that the annotations might also be less accurate in those regions.
2. *Errors in small isolated regions.* This can be observed in Fig. 1f where a small mineralized cartilage region is detected in a bone region.
3. *Large errors on the external border of the bone.* This can be seen on the bottom of the bone region in Fig. 2d. This is explained by the fact that only mineralized cartilage is present on the first micro-CT slices in the 3D image. Bone progressively appears in the imaged structure as we go up in the stack of slices. In the slices where bone starts to appear, the 3D U-net model overestimates the presence of mineralized cartilage. Again, the annotations might have been less accurate in those regions.

Interestingly, the three types of errors appear equally on the three learning strategies presented in this work. The alternative strategies reach an average DSC performance close to 0.85 as well, with one failure case on Set 9 for the third strategy. Since the DSC standard deviation is high compared to the difference of average DSC between the approaches, we cannot identify a significantly better strategy. Hence, a qualitative comparison of the learning strategies is provided in the next sections.

But, before moving to a qualitative comparison, it is worth noting that the DSC only converges to 0.903 (and not 1) on the training set for the first strategy despite the predictions being visually consistent with the ground truth. We might expect that the average DSC measured when comparing annotations provided by different annotators also saturates around 0.9.

Table 2. Comparison of the DSC on the high-resolution test 3D datasets for the different U-net training strategies. Strategy 1: 3D U-net on the low resolution 256 × 256 × 64 3D datasets; Strategy 2: 2D U-net on the low resolution 256 × 256 images; Strategy 3: 3D U-net on the high-resolution 512 × 512 × 16 patches. DSC: Dice Similarity Coefficient. SD: Standard Deviation.

Datasets	Strategy 1	Strategy 2	Strategy 3
Set 1	0.881	0.877	0.858
Set 2	0.889	0.877	0.879
Set 3	0.836	0.850	0.872
Set 4	0.787	0.806	0.812
Set 5	0.868	0.879	0.856
Set 6	0.858	0.857	0.881
Set 7	0.842	0.880	0.816
Set 8	0.878	0.867	0.878
Set 9	0.786	0.860	0.606
Set 10	0.884	0.879	0.899
Set 11	0.861	0.794	0.804
Set 12	0.848	0.781	0.779
Mean ± SD	0.851 ± 0.033	0.851 ± 0.035	0.828 ± 0.076

3.2 Qualitative Impact of the 3D Consistency

Based on Table 2, the 3D U-net model does not show a significant improvement of the DSC performances compared to a 2D U-net approach. The 3D model is however harder to train and test since it is fed with 3D input data, reaching faster the memory bottleneck of the GPU. On the other hand, the 2D approach does not ensure the 3D consistency of the segmentation. This is shown in Fig. 3. Three consecutive slices of Set 2 are shown on it. A discontinuity is present in the segmentation of adjacent slices with the 2D approach, which is not desirable.

3.3 Qualitative Impact of a Reduction in the Spatial Image Resolution

In the first learning strategy, the resolution of the 3D micro-CT datasets is reduced by a factor 4 in every dimension. The ROI prediction is performed on the low resolution datasets as well. In order to report the DSC performance, the low resolution prediction is upsampled to the original resolution. This leads to a rougher edge of the segmented ROI as shown in Fig. 4. Indeed, the low resolution 3D approach is not able to segment accurately the extremity of very irregular regions such as the one shown in Fig. 4c. However, those errors are small compared to the errors presented in Sect. 3.1, as shown in Table 2. Hence, we conclude that reducing the image resolution by a factor 4 preserves most of the information (e.g. the image texture) required to discriminate between the mineralized cartilage and bone.

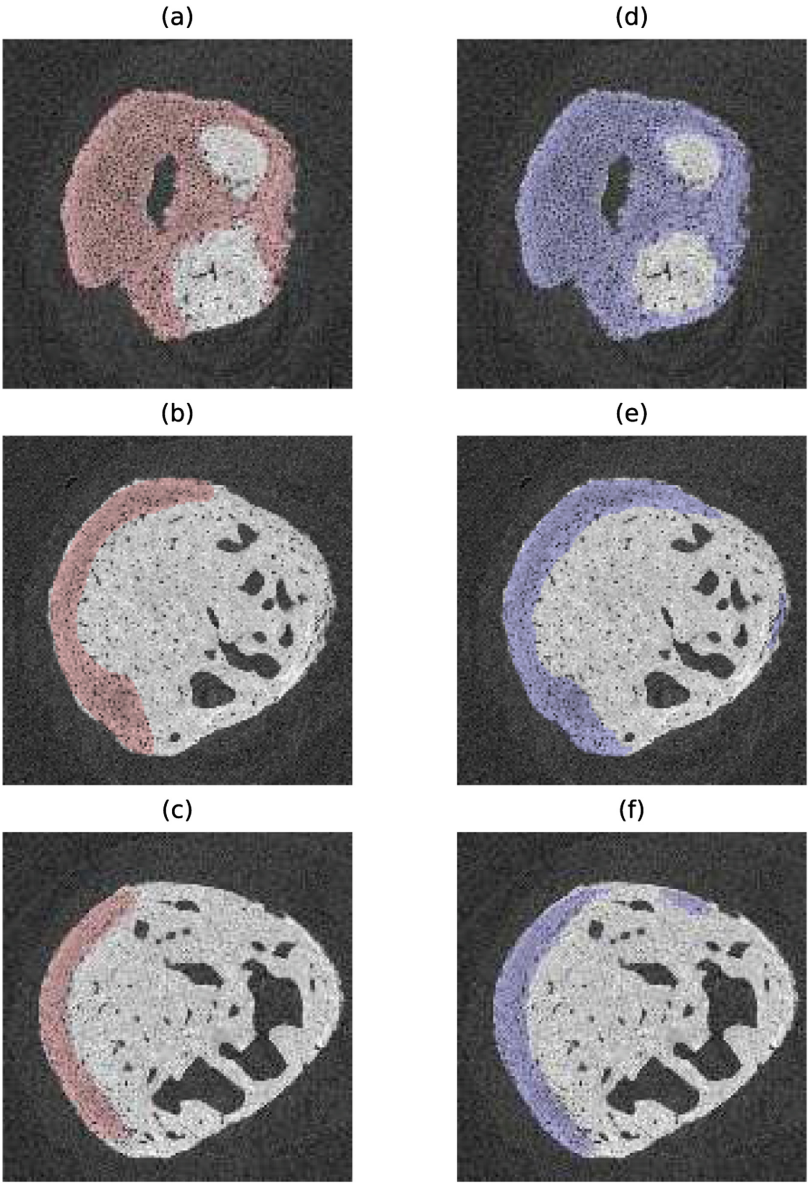

Fig. 1. Comparison of mineralized cartilage segmentations between the ground truth in red (GT) and the prediction of 3D U-net in blue (strategy 1: low resolution) for the slices indices 64, 128, 192 of Set 2.

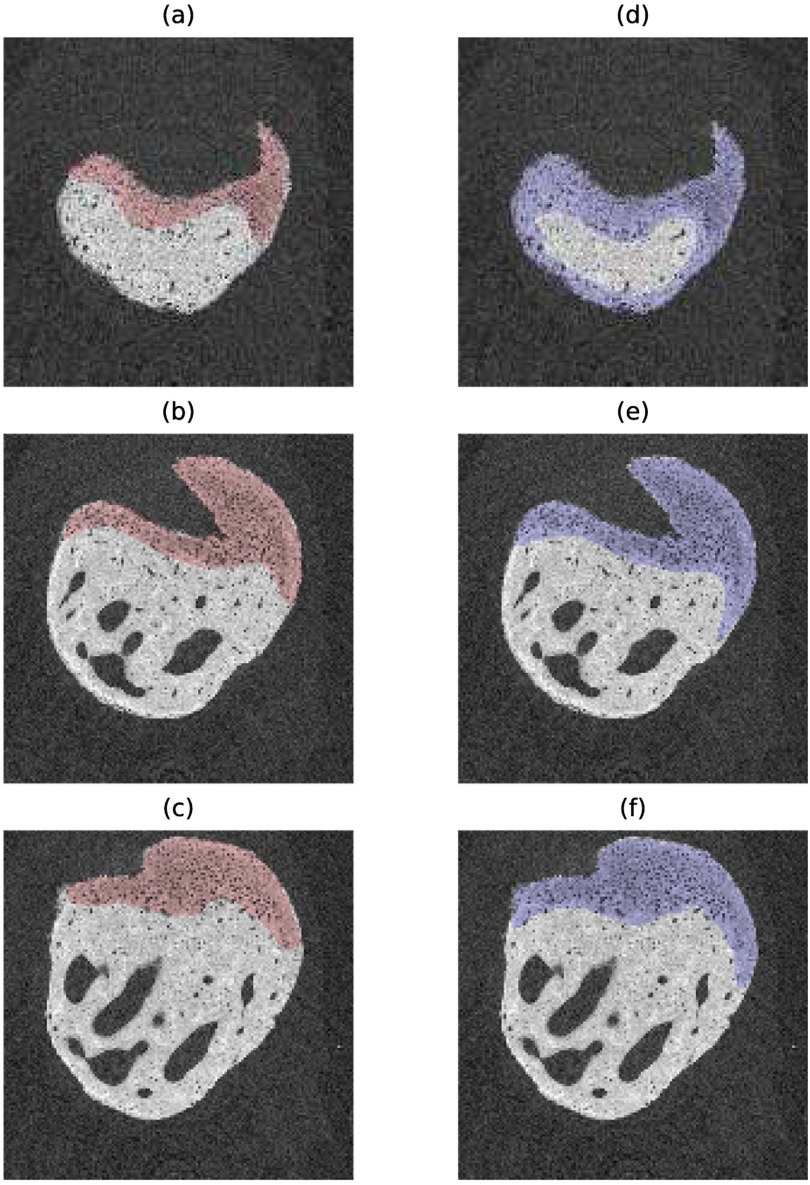

Fig. 2. Comparison of mineralized cartilage segmentations between the ground truth in red (GT) and the prediction of 3D U-net in blue (strategy 1: low resolution) for the slices indices 64, 128, 192 of Set 3.

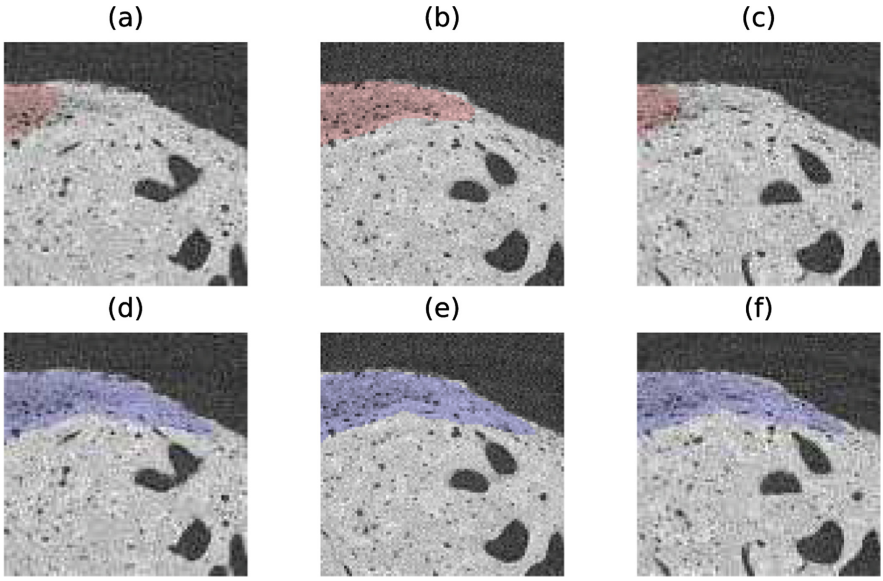

Fig. 3. Comparison of mineralized cartilage segmentations between the 2D U-net in red (strategy 2) and the 3D U-net in blue (strategy 1: low resolution) on a detail of Set 2.

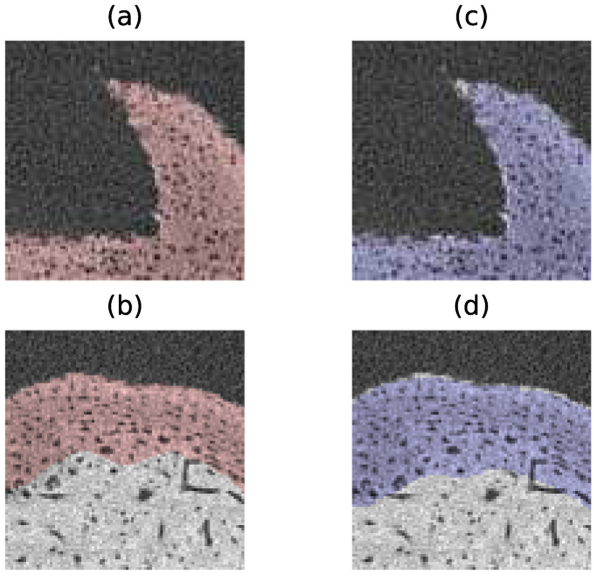

Fig. 4. Comparison of mineralized cartilage segmentations between the patch-based 3D U-net in red (strategy 3) and the low resolution 3D U-net in blue (strategy 1) on a detail of Set 3.

4 Conclusion and Future Perspectives

To our best knowledge, this is the first study presenting a deep learning strategy for automatic segmentation of mineralized cartilage from bone in high-resolution micro-CT images. For this purpose, we use a fully convolutional neural network, i.e. U-net. The average DSC computed on 12 3D micro-CT datasets reaches 0.85 compared to a manually annotated ground truth. The 3D U-net model is compared to the 2D U-net model. The difference in DSC performance is not statistically significant. However, the 3D approach ensures better 3D consistency in the predicted segmentation. The resolution of the micro-CT datasets have been reduced by a factor 4 in every dimensions in order to reduce the training time and to better manage the GPU memory limitations. No major difference in DSC has been observed compared to a 3D U-net trained on full resolution 3D micro-CT patches. Importantly, manual segmentation of mineralized cartilage can be subjected to inter-expert variability. This point has not been considered in the current work and would improve the performance assessment of the proposed methods. Also, the robustness of the deep learning segmentation methods to errors in the annotations should be further investigated since getting high quality annotations is challenging on these kind of datasets.

Acknowledgments. Jean Léger is a Research Fellow of the Fonds de la Recherche Scientifique - FNRS and Christophe De Vleeschouwer is Senior Research Associates with the Belgian F.R.S.-FNRS. Lisa Leyssens acknowledges the Fonds Spéciaux de Recherche (FSR) and the UCLouvain-FSR for funding her PhD. This research project was funded by a research project of the Research Foundation Flanders (FWO; Grant no. G088218N). The high-resolution micro-CT images have been generated at the X-ray computed tomography facilities of the Department of Development and Regeneration of the KU Leuven, financed by the Hercules Foundation (project AKUL 13/47).

References

1. Apostolakos, J., Durant, T.J., Dwyer, C.R., Russell, R.P., Weinreb, J.H., Alaee, F., Beitzel, K., McCarthy, M.B., Cote, M.P., Mazzocca, A.D.: The enthesis: a review of the tendon-to-bone insertion. Muscles Ligaments Tendons J. **4**(3), 333 (2014)
2. Genin, G.M., Kent, A., Birman, V., Wopenka, B., Pasteris, J.D., Marquez, P.J., Thomopoulos, S.: Functional grading of mineral and collagen in the attachment of tendon to bone. Biophys. J. **97**(4), 976–985 (2009)
3. Liu, Y., Thomopoulos, S., Birman, V., Li, J.S., Genin, G.M.: Bi-material attachment through a compliant interfacial system at the tendon-to-bone insertion site. Mech. Mater. **44**, 83–92 (2012)
4. Liu, Y., Birman, V., Chen, C., Thomopoulos, S., Genin, G.M.: Mechanisms of bimaterial attachment at the interface of tendon to bone. J. Eng. Mater. Technol. **133**(1), 011006 (2011)

5. Kerckhofs, G., Sainz, J., Wevers, M., Van de Putte, T., Schrooten, J.: Contrast-enhanced nanofocus computed tomography images the cartilage subtissue architecture in three dimensions. Eur. Cells Mater. **25**, 179–89 (2013)
6. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 424–432. Springer (2016)
7. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
8. Brion, E., Léger, J., Javaid, U., Lee, J., De Vleeschouwer, C., Macq, B.: Using planning CTs to enhance CNN-based bladder segmentation on cone beam CT. In: Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling, vol. 10951, p. 109511M. International Society for Optics and Photonics (2019)
9. Fripp, J., Crozier, S., Warfield, S.K., Ourselin, S.: Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. IEEE Trans. Med. Imaging **29**(1), 55 (2010)
10. Tamez-Pena, J.G., Farber, J., Gonzalez, P.C., Schreyer, E., Schneider, E., Totterman, S.: Unsupervised segmentation and quantification of anatomical knee features: data from the osteoarthritis initiative. IEEE Trans. Biomed. Eng. **59**(4), 1177–1186 (2012)
11. Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M.: Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: International Conference on Medical Image Computing and Computer-assisted Intervention, pp. 246–253. Springer (2013)
12. Liu, F., Zhou, Z., Jang, H., Samsonov, A., Zhao, G., Kijowski, R.: Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magn. Reson. Med. **79**(4), 2379–2391 (2018)
13. Norman, B., Pedoia, V., Majumdar, S.: 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 **288**(1), 177–185 (2018)
14. Zhou, Z., Zhao, G., Kijowski, R., Liu, F.: Deep convolutional neural network for segmentation of knee joint anatomy. Magn. Reson. Med. **80**(6), 2759–2770 (2018)
15. Ambellan, F., Tack, A., Ehlke, M., Zachow, S.: Automated segmentation of knee bone and cartilage combining statistical shape knowledge and convolutional neural networks: data from the osteoarthritis initiative. Med. Image Anal. **52**, 109–118 (2019)
16. Raj, A., Vishwanathan, S., Ajani, B., Krishnan, K., Agarwal, H.: Automatic knee cartilage segmentation using fully volumetric convolutional neural networks for evaluation of osteoarthritis. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 851–854. IEEE (2018)
17. Tabrizi, P.R., Zoroofi, R.A., Yokota, F., Tamura, S., Nishii, T., Sato, Y.: Acetabular cartilage segmentation in CT arthrography based on a bone-normalized probabilistic atlas. Int. J. Comput. Assist. Radiol. Surg. **10**(4), 433–446 (2015)
18. Myller, K.A., Honkanen, J.T., Jurvelin, J.S., Saarakkala, S., Töyräs, J., Väänänen, S.P.: Method for segmentation of knee articular cartilages based on contrast-enhanced CT images. Ann. Biomed. Eng. **46**(11), 1756–1767 (2018)

19. Frondelius, T., Tiulpin, A., Lehenkari, P., Nieminen, H., Saarakkala, S.: Fully automatic deep learning based segmentation of bone-cartilage interface from micro-CT images of human osteochondral samples. Osteoarthritis Cartilage **26**, S469 (2018)
20. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234–241. Springer (2015)