

## Baker's Cyst Classification Using Random Forests

Adam Ciszkiewicz, Grzegorz Milewski  
Institute of Applied Mechanics  
Cracow University of Technology,  
al. Jana Pawła II 37, 31-864 Cracow, Poland  
Emails: acisz@poczta.fm,  
milewski@mech.pk.edu.pl

Jacek Lorkowski  
Department of Orthopaedics and Traumatology  
Central Clinical Hospital  
of the MSWiA in Warsaw  
Wołoska 137, 02-507 Warsaw, Poland  
Email: jacek.lorkowski@gmail.com

**Abstract**—In this paper, a classification procedure for Baker's cysts was proposed. The procedure contained two subprocedures: the image preprocessing (dual thresholding, labeling, feature extraction) and the classification (Random Forests, cross validation). In total, five features were required to classify the cysts. These geometric features represented the location, the area and the convexity of the cyst. The procedure was proven effective on a set 436 varied MRI images. The set contained 68 images with cysts ready for aspiration and was oversampled with the SMOTE approach. The proposed method operates on 2D MRI images. This reduces the time of diagnosis and, with the ever increasing demand for MRI scanners, is justified economically. The method can be employed in systems for autonomous and semi-autonomous Baker's cyst aspiration or as a standalone package for MRI images annotation. Furthermore, it can be also extended to other fluid-based medical conditions in the knee.

### I. INTRODUCTION

WITH the ageing of the world population and the ever present health sector shortages [1], automation of surgical procedures is increasingly more common and important. For some of the simpler medical conditions, it is possible to develop robotic systems to treat them nearly autonomously. This in turn offloads the medical staff – their time and experience can be devoted to difficult and complicated surgical procedures.

A Baker's cyst is a very common medical condition (see Fig. 1). The cyst is a synovial capsule filled with fluid, which often occurs when the knee is in an inflammatory state. This condition is not dangerous, but it can affect the

patient's quality of life through pain and reduced range of motion of the knee. The safest and easiest method of treatment for this condition is aspiration [2]. The aspiration is usually preceded by Magnetic Resonance Imaging (MRI) [3]. Often, full 3D scans are obtained, as they can be used to precisely diagnose the cyst and chose the best approach for aspiration. Despite its simplicity, fully autonomous aspiration of Baker's cyst remains largely unexplored in the literature. To automate this procedure, it is necessary to propose a classification procedure, capable of determining whether the knee contains a Baker's cyst that is ready for aspiration.

Two major approaches to medical image classification can be distinguished. In the first one, the learning algorithm is supplied a raw image and it learns meaningful features from the image automatically [4]. This approach is intuitive but requires large training sets and time-consuming training to provide good results. Such datasets can only be obtained through scientific collaborations, which are often focused on pressing medical issues. In case of simpler medical conditions, the second approach to classification can be employed. This method involves feature extraction coupled with image processing [5]. As the features are defined by the user, smaller training sets can be used. To extract the features it is often necessary to segment the image first. While there are many methods available for the medical image segmentation [6], the thresholding remains one of the more popular approaches [7]. The threshold can be set manually by the user or automatically with one of the available methods [8], [9]. Both approaches were utilized in this study.

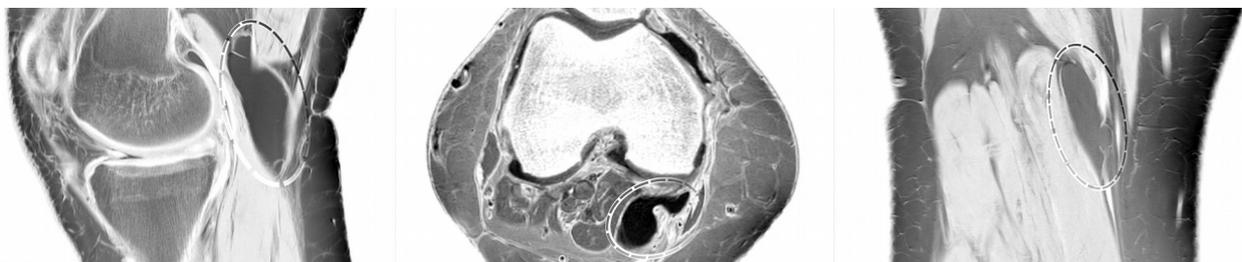


Fig 1. The Baker's cyst.

In a broader context, a classification procedure can be seen as a module in systems for autonomous and semi-autonomous surgery, in which it is used for diagnosis. The other problems in this area include: physical simulation of soft tissues and joints, medical tool path planning, surgery planning and surgery optimization [10]–[15].

In this paper, a procedure for Baker's cyst classification was proposed. The procedure was based on Random Forests and operated on 2D scans from Magnetic Resonance Imaging (MRI). The scans were preprocessed using custom, dual thresholding and labeling. Then, five geometric features were extracted from the preprocessed images. These features were then used to train the Random Forest (with cross-validation). The procedure was repeated 40 times to ensure that the obtained results were independent of the random number generator.

## II. METHOD

### A. The input dataset

The input dataset contained 12 spatial MRI sets of the knee. In total, the number of 2D MRI slices was 436. Out of them only 68 contained a Baker's cyst that was ready for aspiration. These 68 images were selected by an experienced orthopedic surgeon. The disproportion between the classes is very common in medical classification problems. Typical machine learning algorithms don't perform well with such datasets. Therefore, it is necessary to balance them. In this study, the input dataset was balanced with the Synthetic Minority Over-sampling Technique (SMOTE) [16] using Imbalanced-learn [17].

The proposed procedure was tested on three different MRI sequences: PDW SPAIR (Proton Density Weighted Spectral Attenuated Inversion Recovery), STIR (Short Tau Inversion Recovery) and PDW FatSat (Proton Density Weighted Fat Saturation) MRI sequences. In all of these sequences, the fluid-based structures appear hyper-intense, as seen in Fig. 2. The images were imported into Python using Pydicom [18].

The Baker's cysts are typically diagnosed with 3D MRI scans. The idea to use 2D slices instead of full 3D sets was partially inspired by the FAST USG, often employed in abdominal cavity diagnosis. In FAST USG the patient is diagnosed with only a few sweeps of the transducer. The diagnosis takes between 10 and 20 seconds, which is less expensive than a full sweep. These advantages also apply to MRI. With the ever growing need for MRI-based diagnosis, the

availability of MRI scanners is a significant factor. Furthermore, some patients can't undergo full MRI scanning and others may feel uncomfortable during the procedure. When using singular 2D MRI scans for diagnosis, as in the proposed procedure, these issues are no a longer a concern.

### B. Image processing

The initial threshold for the image segmentation was obtained using the method proposed in [9]. This threshold was then modified as follows:

$$thr_{fin} = mfn[\mathbf{I}(x, y) * (\mathbf{I}(x, y) > thr)], \quad (1)$$

where:  $thr_{fin}$  – the final threshold,  $thr$  – the threshold computed with the method presented in [9],  $\mathbf{I}(x, y)$  – the input image,  $mfn()$  – a function that computes the mean value of a matrix using only the nonzero elements.

In the next step the image was labeled with a 3 by 3 pixel mask. After the labeling, the largest object in the segmented image was selected. This object (see Fig. 3b) was assumed to be the cyst candidate. In the third step, the boundary of the knee was obtained (see Fig. 3c) using the following threshold:

$$thr_b = k \max[\mathbf{I}(x, y)] + (1 - k) \min[\mathbf{I}(x, y)], \quad (2)$$

where:  $k$  – the thresholding parameter (here:  $k = 0.07$ ). Finally, the obtained bounding box of the knee was applied to the image of the segmented object (see Fig. 3d).

### C. Features

In this study, the following 5 features were used in the classification procedure:

a)  $object_{area}$  – a typical feature used in medical classification procedures, which represented the area of the segmented object [ $mm^2$ ],

b)  $circularity$  – a dimensionless measure of how circular the segmented object was; computed as a ratio between the area of the Largest Empty Circle (LEC) inside the object  $LEC_{area}$  and the object area  $object_{area}$ :

$$circularity = LEC_{area} / object_{area}. \quad (3)$$

The LEC (see Fig. 4b) was obtained using the method presented in [19],

c)  $convexity$  – a relative feature that measured the convexity of the segmented object; computed as a ratio between the object area  $object_{area}$  and the area of the convex hull of the object  $convexhull_{area}$  (see Fig. 4a):

$$convexity = object_{area} / convexhull_{area}. \quad (4)$$

The convex hull of the object was computed using Scipy [20],

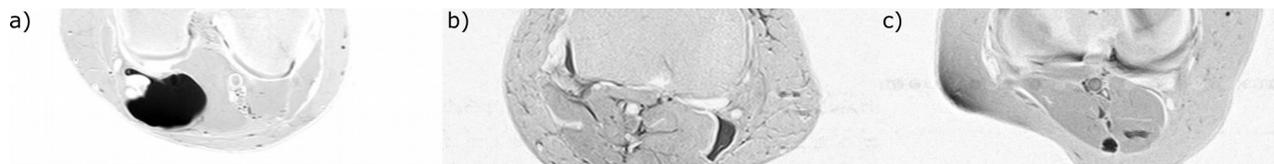


Fig. 2. The supported MRI sequences: a) PDW SPAIR, b) STIR, c) PDW FatSat.

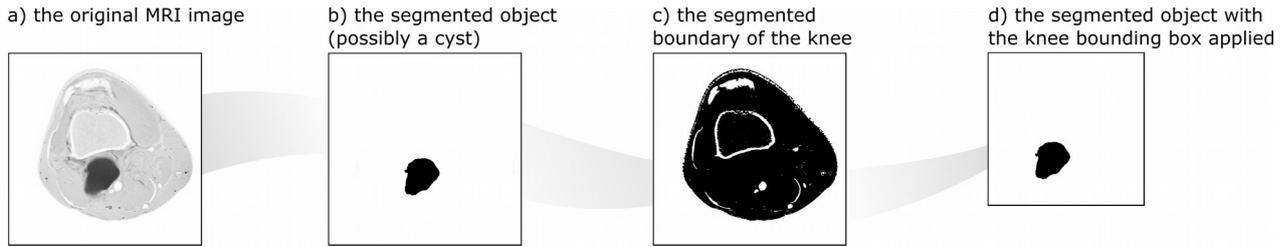


Fig 3. The image processing.

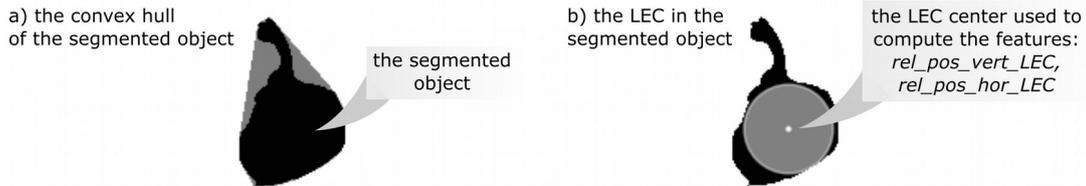


Fig 4. The convex hull and the LEC in the segmented image.

d)  $rel\_pos\_vert\_LEC$  – the relative vertical position of the LEC center (see Fig. 4b),

e)  $rel\_pos\_hor\_LEC$  – the relative horizontal position of the LEC center.

The first feature separated the cysts from the veins and the arteries – the cysts ready for aspiration often had larger areas. The second and the third feature helped distinguish cases with segmentation errors (multiple, small and connected objects or images with no hyper-intense objects). Furthermore, the second and the third feature were computed by dividing the area of the LEC and the convex hull respectively by the object\_area. This made them independent of the first feature.

The fourth and the fifth feature differentiated the cysts from other fluid-related conditions in the knee, such as knee effusions – the cysts usually occur in the posterior side of the knee. With the proposed skin segmentation procedure, the LEC center was computed with regards to the actual boundary of the knee – not the image (see Fig. 3d). This made the procedure more general and reliable.

#### D. The classifier

Decision trees are among the most popular classification algorithms in machine learning. They are easy to implement and quick to train. Despite these advantages, the trees tend to overfit data. This is a serious drawback, especially when working with limited datasets. A natural extension to this approach, which addresses this flaw, is Random Forest [21]. In this algorithm several decision trees are trained on different subsets of the training dataset and the final classification is based on majority vote. It is worth mentioning that, unlike many classification algorithms, Random Forests do not require data preconditioning. This means that raw feature values can be used to train them and classify new samples.

In this study, Random Forest contained 10 decision trees, based on Gini impurity. The depth of the trees was not limited. The classifier was implemented using Scikit-learn [22]. The training and the testing were performed with a 4-fold cross validation, while the performance of the classifier was measured with the following indicators:

$$\begin{aligned} Sen &= TP / (TP + FN), \\ Spec &= TN / (TN + FP), \\ F_1 &= 2 * TP / (2 * TP + FN + FP), \end{aligned} \quad (5)$$

where:  $Sen$  – the sensitivity,  $Spec$  – the specificity,  $F_1$  – the  $F_1$  score (also referred to as balanced  $F$ -score),  $TP$  – the number of true positive cases (the knee contains a cyst ready for aspiration and is classified for aspiration),  $FP$  – the number of false positive cases (the knee does not contain a cyst ready for aspiration and is classified for aspiration),  $TN$  – the number of true negative cases (the knee does not contain a cyst ready for aspiration and is not classified for aspiration),  $FN$  – the number of false negative cases (the knee contains a cyst ready for aspiration and is not classified for aspiration).

### III. RESULTS AND DISCUSSION

The training and test procedures were repeated 40 times to factor in the effects of shuffling and oversampling. The average and the best (based on the  $F_1$ -score) results over the 40 runs were summarized in Table 1.

As seen in Table 1, the mean values of the performance indicators were higher than 95.0%. Furthermore, the best run achieved  $F_1$  of 99.4 % and specificity of 100.0 %. This proves that the proposed method is capable of classifying the cysts for aspiration based on 2D MRI images. It is worth mentioning, that this classification problem would be easier with 3D MRI images. In three dimensions, the cyst can be easily distinguished from veins and arteries based on its spherical shape. Nevertheless, 3D MRI scanning can be time

TABLE I.  
THE PERFORMANCE INDICATORS BASED ON THE 40 RUNS

	<i>TP</i>	<i>FP</i>	<i>TN</i>	<i>FN</i>	<i>Sen</i> [%]	<i>Spec</i> [%]	$F_1$ [%]
<i>avg</i>	88.6± 3.0	2.0± 3.0	94.0± 2.8	4.1± 3.8	95.6± 4.2	97.8 ± 3.4	96.7± 2.6
<i>best</i>	89	0	94	1	98.9	100.0	99.4

consuming. Furthermore, the demand for MRI scanners exceeds the supply in most medical centers in the world. Therefore, good performance using 2D slices can be seen as an advantage of the proposed procedure.

Currently, the procedure analyzes only the largest segmented object in the knee (these objects were classified by the surgeon and used to train/test the Random Forest). Nevertheless, in some rare cases the cyst may not be the largest segmented object. This issue can be solved by iteratively checking and classifying all of the segmented objects.

#### IV. CONCLUSION

In this paper, a classification procedure for Baker's cyst was proposed. The procedure was composed of two subroutines: the image preprocessing (coupled with feature extraction) and classification using Random Forests. The procedure was proven effective on a set 468 varied MRI images. The images were obtained using three, different MRI sequences. Good performance using 2D slices can be seen as an advantage of the proposed procedure. This reduces the time of diagnosis, improves patient's comfort and, with the ever increasing demand for MRI scanners, is justified economically. The method can be employed in systems for autonomous and semi-autonomous Baker's cyst aspiration or as a standalone package for MRI images annotation. Furthermore, it can be also extended to other fluid-based medical conditions in the knee.

#### REFERENCES

- [1] X. Scheil-Adlung, T. Behrendt, and L. Wong, "Health sector employment: A tracer indicator for universal health coverage in national Social Protection Floors," *Hum. Resour. Health*, vol. 13, no. 1, pp. 1–8, 2015, <http://dx.doi.org/10.1186/s12960-015-0056-9>
- [2] S. Jamshed and L. M. Snyder, "An Intact Dissecting Baker's Cyst Mimicking Recurrent Deep Vein Thrombosis," *J. Investig. Med. High Impact Case Reports*, vol. 4, no. 2, p. 2324709616650703, Apr. 2016, <http://dx.doi.org/10.1177/2324709616650703>
- [3] T. J. Frush and F. R. Noyes, "Baker's Cyst: Diagnostic and Surgical Considerations.," *Sports Health*, vol. 7, no. 4, pp. 359–65, Jul. 2015, <http://dx.doi.org/10.1177/1941738113520130>
- [4] W. Sun, T.-L. Tseng, J. Zhang, and W. Qian, "Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data," *Comput. Med. Imaging Graph.*, vol. 57, pp. 4–9, 2017, <http://dx.doi.org/10.1016/j.compmedimag.2016.07.004>
- [5] P. Zarychta, P. Badura, and E. Pietka, "Comparative analysis of selected classifiers in posterior cruciate ligaments computer aided diagnosis," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 65, no. 1, pp. 63–70, 2017, <http://dx.doi.org/10.1515/bpasts-2017-0008>
- [6] D. J. Withey and Z. J. Koles, "Three Generations of Medical Image Segmentation: Methods and Available Software," *Int. J. Bioelectromagn.*, vol. 9, no. 2, pp. 67–68, 2007.
- [7] T. Markiewicz et al., "Thresholding techniques for segmentation of atherosclerotic plaque and lumen areas in vascular arteries," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 63, no. 1, pp. 269–280, 2015, <http://dx.doi.org/10.1515/bpasts-2015-0031>
- [8] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979, <http://dx.doi.org/10.1109/TSMC.1979.4310076>
- [9] J. C. Jui-Cheng Yen, F. J. Fu-Juay Chang, and S. Shyang Chang, "A new criterion for automatic multilevel thresholding," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 370–378, Mar. 1995, <http://dx.doi.org/10.1109/83.366472>
- [10] O. A. Pappalardo et al., "Mass-spring models for the simulation of mitral valve function: Looking for a trade-off between reliability and time-efficiency," *Med. Eng. Phys.*, vol. 47, pp. 93–104, 2017, <http://dx.doi.org/10.1016/j.medengphy.2017.07.001>
- [11] N. Chentanez et al., "Interactive Simulation of Surgical Needle Insertion and Steering," *ACM Trans. Graph.*, vol. 28, no. 3, 2009, <http://dx.doi.org/10.1145/1531326.1531394>
- [12] G. J. Vrooijink, M. Abayazid, S. Patil, R. Alterovitz, and S. Misra, "Needle path planning and steering in a three-dimensional non-static environment using two-dimensional ultrasound images," *Int. J. Rob. Res.*, vol. 33, no. 10, pp. 1361–1374, Sep. 2014, <http://dx.doi.org/10.1177/0278364914526627>
- [13] K. C. Assi, S. Grenier, S. Parent, H. Labelle, and F. Cheriet, "A physically based trunk soft tissue modeling for scoliosis surgery planning systems.," *Comput. Med. Imaging Graph.*, vol. 40, pp. 217–28, Mar. 2015, <http://dx.doi.org/10.1016/j.compmedimag.2014.11.002>
- [14] A. Ciszakiewicz, J. Lorkowski, and G. Milewski, "A novel planning solution for semi-autonomous aspiration of Baker's cysts," *Int. J. Med. Robot.*, p. e1882, 2018, <http://dx.doi.org/10.1002/rcs.1882>
- [15] A. Ciszakiewicz and G. Milewski, "Path planning for minimally-invasive knee surgery using a hybrid optimization procedure," *Comput. Methods Biomech. Biomed. Engin.*, vol. 21, no. 1, 2018, <http://dx.doi.org/10.1080/10255842.2017.1423289>
- [16] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002, <http://dx.doi.org/10.1613/jair.953>
- [17] G. Lemaitre, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning," *J. Mach. Learn. Res.*, vol. 18, no. 17, pp. 1–5, 2017.
- [18] D. Mason, "SU-E-T-33: Pydicom: An Open Source DICOM Library," *Med. Phys.*, vol. 38, no. 6, p. 3493, 2011, <http://dx.doi.org/10.1118/1.3611983>
- [19] G. T. Toussaint, "Computing largest empty circles with location constraints," *Int. J. Comput. Inf. Sci.*, vol. 12, no. 5, pp. 347–358, 1983.
- [20] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Comput. Sci. Eng.*, vol. 13, no. 2, pp. 22–30, Mar. 2011, <http://dx.doi.org/10.1109/MCSE.2011.37>
- [21] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, <http://dx.doi.org/10.1023/A:1010933404324>
- [22] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2012.