

Automatic Brain Localization in Fetal MRI Using Superpixel Graphs

Amir Alansary¹(✉), Matthew Lee¹, Kevin Keraudren¹, Bernhard Kainz^{1,2},
Christina Malamateniou², Mary Rutherford², Joseph V. Hajnal²,
Ben Glocker¹, and Daniel Rueckert¹

¹ Department of Computing, Imperial College London, London, UK
`{a.alansary14,matthew.lee13,kevin.keraudren10,
b.kainz,b.glocker,d.rueckert}@imperial.ac.uk`

² Division of Imaging Sciences, King's College London, London, UK
`{christina.malamateniou,mary.rutherford,jo.hajnal}@kcl.ac.uk`

Abstract. Fetal MRI is emerging as an effective, non-invasive tool in prenatal diagnosis and pregnancy follow-up. However, there is a significant variability of the position and orientation of the fetus in the MR images. This makes these images more difficult to analyze and interpret compared to standard adult MR imaging, which standardized anatomical imaging aligned planes. We address this issue by automatic localization of the fetal anatomy, in particular, the brain which is a structure of interest for many fetal MRI studies. We first extract superpixels followed by the computation of a histogram of features for each superpixel using bag of words based on dense scale invariant feature transform (DSIFT) descriptors. We construct a graph of superpixels and train a random forest classifier to distinguish between brain and non-brain superpixels. The localization framework has been tested on 55 MR datasets at gestational ages between 20–38 weeks. The proposed method was evaluated using 5-fold cross validation achieving a 94.55 % brain detection accuracy rate.

1 Introduction

Fetal magnetic resonance imaging (MRI) has significantly improved in the last two decades, and is emerging as a novel, non-invasive tool for diagnosis and planing of surgical interventions. It provides higher contrast and larger field-of-view than ultrasound. Thus, it provides better structural information of the different fetal organs such as the brain, spine and body. Fetal brain localization is important for assessing the fetal brain development and maturation. It is also the primary step for most of the current automatic motion correction techniques for fetal MRI [10]. Recently, fetal brain detection has been used as a landmark to extract the other fetal organs [11]. Problems that hinder the design of automated image analysis tools for fetal MRI usually arise from: (a) the high variability in shape, size, orientation, and anatomical configuration of the fetus; (b) intensity non-uniformities (bias artifacts); (c) partial volume effects; and (d) motion artifacts caused by the unconstrained fetal motion (see Fig. 1).

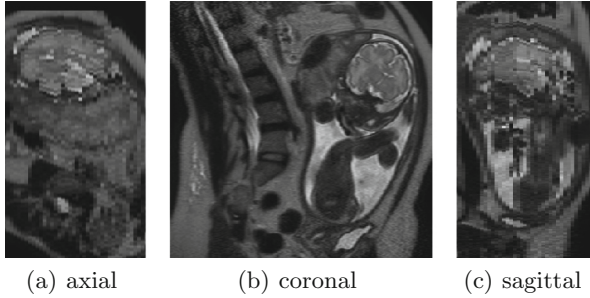


Fig. 1. Three orthogonal cutting planes through a stack of fetal MRI images. The quality of the in-plane (coronal) slices is not affected by motion, however, there are inter-slice artifacts appear in the out-of-plane views (axial and sagittal).

Related work: Fetal MRI is an emerging field of research, with little work focused on fully automatic processing of these datasets. In [2], 3D template matching is used to detect the eyes, enabling a subsequent 2D/3D graph-cut segmentation that extracts the brain. This approach is based on 3D templates and lacks the flexibility necessary to deal with motion artifacts as well as fetal abnormalities. The methods proposed in [9] and [12] address the variability of fetal MRI through machine learning. In [9], a Random Forest (RF) classifier first distinguishes between maternal and fetal tissues before classifying different tissues of the fetal head, while [12] combines prior knowledge of the fetal size with maximally stable extremal regions (MSER) detection and a bag-of-words model.

Contribution: In this paper we propose a fully-automated framework for localizing the fetal brain in fetal MRI scans. Rather than working on individual pixels we make use of superpixels for a faster and more efficient detection algorithm. Because of the nature of superpixels that most likely represents the rigid regions in the image, using superpixels neighbors instead of pixel neighbors can reduce the effect of motion artifacts. Therefore, we have developed a new superpixel graphical model based on both spatial and intensity distances in 3D. The proposed localization framework achieves 94.55 % accuracy for the brain detection and 98.18 % prediction accuracy of the center of the brain. The proposed approach does not require landmarks as in [9] or prior information such as the gestational age of the fetus as in [12].

2 Method

The proposed approach for the automatic localization of the fetal brain consists of four main steps as shown in Fig. 2. The input data of our system are 3D fetal MRI datasets. The first step is to decompose each 2D slice into $\{N_1, \dots, N_s\}$ superpixels in order to minimize the local redundancy in the input data. By clustering and constructing a single descriptor for each superpixel we reduce the impact of noise on each descriptor whilst preserving homogeneous regions that

are likely belong to the same anatomical region. The second step is to calculate image descriptors for each pixel and then aggregate them into one histogram h_i for each i -superpixel. The third step is to build superpixel graphs based on each superpixel's neighbors. Then each superpixel's histogram is normalized with its neighbors in the graphical model. During the fourth step, we use a random forest to generate a probability map of the brain for every superpixel. Finally, this probability map is refined using another auto-context classifier followed by selecting the largest 3D component.

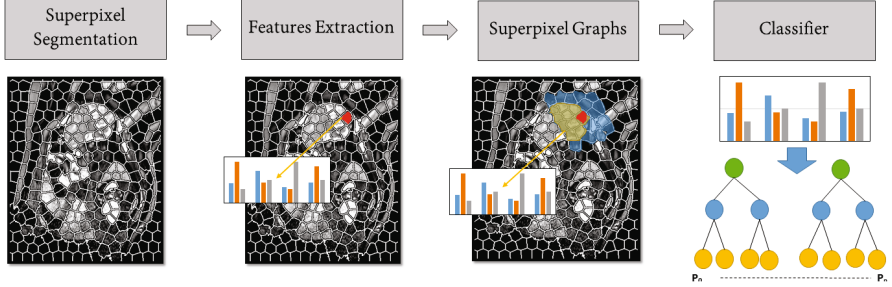


Fig. 2. The automatic localization of the fetal brain framework.

Superpixels: Superpixels are a popular unsupervised image segmentation technique that clusters image pixels into groups of pixels based on the correlation between each pixel and its neighbors. In the literature [1], many superpixel techniques have been shown to produce ‘good’ segmentation results. However, which properties of superpixels are important depends on the application. Medical image analysis can be computationally expensive when compared to normal image analysis due to the size of the data, namely because MRI scans are 3D volumes as opposed to 2D images.

Most of the current superpixel segmentation techniques have been proposed for 2D images. In our work, we have chosen the simple linear iterative clustering (SLIC) technique [1], which is fast to compute while achieving a good segmentation quality (as shown in Fig. 3) with lower computational cost so that the method scales well when processing the many slices of a volume. SLIC segments pixels into compact and nearly uniform superpixels. Superpixels are applied in 2D (not 3D) because of the fetal motion that results in 2D misaligned slices. Choosing the right number of superpixels for each 2D slice is challenging. Thus, we have modified the ad-hoc heuristics proposed by [8] to optimize superpixels for fetal MRI. We have weighted the rule with a constant factor k , where $k \in \mathbb{R}_{>0}$ and is chosen depending on n , the total number of pixels in the 2D input image. Thus, the total number of superpixels $s = |\{N_1, \dots, N_s\}|$ in a 2D slice has been calculated:

$$s \approx k \cdot \sqrt{(n/2)}. \quad (1)$$

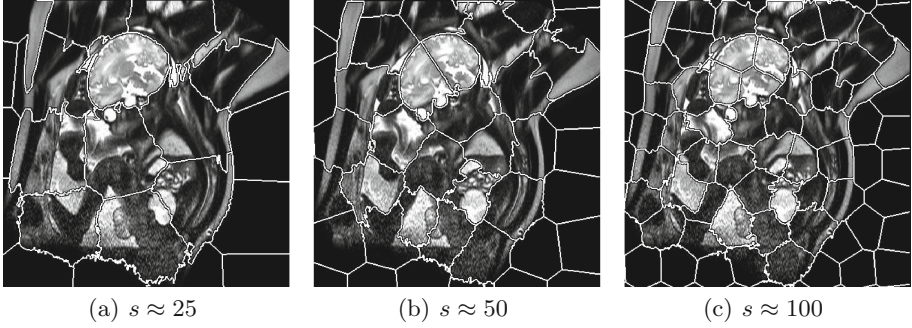


Fig. 3. A cropped image of a 2D fetal MRI scan segmented at (a) 25-, (b) 50-, and (c) 100-superpixels.

Image descriptors: We build a bag of features using dense scale invariant feature transform (DSIFT) [16] descriptors. This is done by first computing SIFT [13] descriptors for each pixel in every 2D image in the training set at a fixed scale and orientation. A k -means clustering is then performed on these descriptors and their centers are used to form a dictionary of k words. When collecting DSIFT descriptors from pixels we can then find their closest matching word from this dictionary, aggregating the frequency of words in each superpixel N_i into one histogram h_i with k -bins. The k -dimensional histogram acts as the feature descriptor for each superpixel. However, the descriptor is constructed in such a way that only contains local information about the superpixel itself. This leads to a loss of the large-scale image context. Also, due to the nature of the superpixels, their histograms of features tend to be sparse. Most of the DSIFT descriptors within a superpixel are likely to be mapped to the same word.

Superpixel graphs: To overcome the problem of sparse descriptors for superpixels, we construct superpixel graphs using the distance between the centroids of superpixels as edges [7]. These edges are weighted based on both spatial and geodesic distances. We first compute the centroid c_i for each N_i superpixel. Then we identify r neighbors based on the similarity score between two superpixels N_i and N_o with centroids c_i and c_o . The similarity between two superpixels is defined:

$$f_r(N_i, N_o) = 1 - \frac{d(c_i, c_o)}{D}, \quad (2)$$

Here $d(\cdot, \cdot)$ is the Euclidean distance, and D is the length of the diagonal of the 2D image. This normalizes the score to be between $[0, 1]$. The closer $f_r(c_i, c_o)$ is to 1, the smaller is the distance between these two centroids. We next assign weights w_j for the extracted r neighbors candidates based on the geodesic distances between their centroids. The geodesic distance is estimated using:

$$d_r(N_i, N_o) = \sum_{j=1}^m |I(p_j) - I(p_{j-1})|, \quad (3)$$

Here $I(p_j)$ is the intensity of the pixel p_j , and m equals to the number of pixels located on the straight line between c_i and c_o . Next, the weights w_j are calculated by normalizing d_r , so that $w_j = 1$ when d_r is the lowest and $w_j = 0$ when d_r is the highest. Finally, the histograms of features for the extracted r superpixels are aggregated based on the calculated weights, and normalized using:

$$\tilde{h}_i = \frac{\sum_{j=0}^r w_j h_j}{\left\| \sum_{j=0}^r w_j h_j \right\|_{\ell_1}} \quad (4)$$

Here $h_0 = h_i$ or the histogram of the current superpixel in consideration. Using graphs of superpixels enables the proposed localization method to overcome the motion artifacts between the 2D slices by extracting the superpixel neighbors in 3D, which would be more difficult using graphs of individual pixels. This is because of the nature of superpixels that most likely represent rigid regions, see Fig. 4. By selecting both spatial and intensity neighbors, we increase the features for each extracted histogram instead of using sparse features. Consequently, we extend the features used in the machine learning to include the image context instead of using only the local information. Figure 4 shows the proposed graphical model in both 2D and 3D.

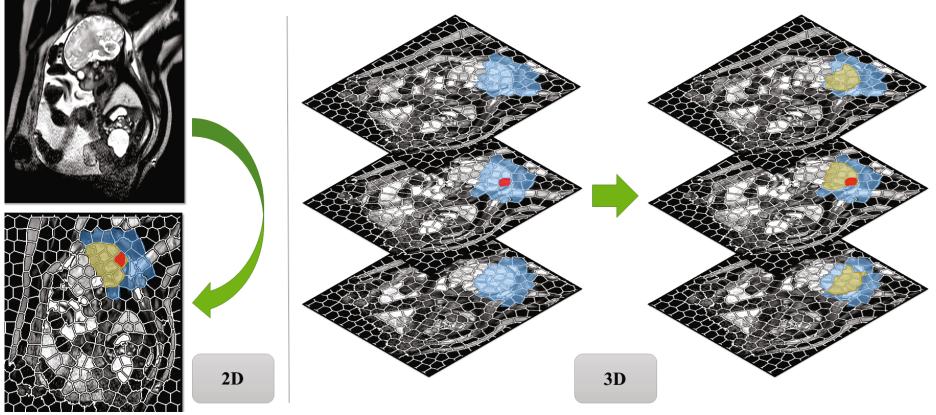


Fig. 4. The proposed superpixel graphical model in 2D and 3D. The red superpixel represents the current superpixel in consideration, the r -spatial neighbors are colored in blue, and the neighbors with higher weights (w_j) are colored in green (Color figure online).

Classification: We use the normalized feature vectors to train a two class random forest (RF) [4] to classify each superpixel as brain or non-brain. Our dataset has pixel-wise labels so we assign each superpixel N_i the class label l_i that corresponds to the label with the highest frequency inside the superpixel. After the initial classification we obtain a probability map and use this to train another second random forest along with 10 % of the most important features used for training the first classifier. This produces an auto-context classifier [15] that can increase the classification accuracy. The output volume is then filtered by finding the largest 3D component, which is the brain mask in this case, followed by a convex-hull extraction [3] to obtain a clean homogeneous segmentation.

3 Evaluation and Results

Data: The proposed framework has been tested on 55 fetal scans at gestational age between 20–38 weeks. Thirty subjects of these datasets are from normal fetuses and 25 datasets are from fetuses with intrauterine fetal growth restriction (IUGR). The data was acquired with a 1.5T Philips MRI system using single shot fast spin echo (ssFSE) sequences with voxel size $0.8398 \times 0.8398 \times 4 \text{ mm}^3$. Ground truth labels were obtained by manual segmentation of the brain performed by expert observers.

Implementation: We perform mean and standard deviation normalization on the input scan intensities as a preprocessing stage for our proposed approach. We have adjusted the SLIC superpixel extraction used in [1] for generating superpixels that are optimized for fetal MRI data. The vlfeat library [16] was used for generating DSIFT features and the scikit-learn library [14] was used for the random forest classifier. The code was implemented using python and MatLab with the mex-c environment. We use $k = 5$ to determine the number of superpixels s . The number of neighbors selected for superpixel graphs r were set to 25 calculated for each xy -plane in three slices. In order to balance the positive and negative training samples for the classifier, we choose to restrict training to superpixels generated from a cropped volume around the brain by adding 25 % of the maximum brain diameter in the xy -plane. The prior knowledge that brains appear brighter in T2-scans allowed us to suppress some of the background pixels by thresholding any pixel less than 10 % of the maximum intensity value of the whole subject. All these parameters are chosen by experiment on a smaller test dataset.

Results: A 5-fold cross validation was used for evaluating our approach (11 test patients 44 training patients per fold). The random forest classifier achieved an average accuracy score of 96.17 % per a superpixel basis. We defined the detection accuracy, the extracted mask covering at least 70 % of the brain, similar to the definition presented in [9] but calculated for the whole 3D brain for simplification. The prediction accuracy of the centers of the segmented brains are measured by calculating the percentage of centers that lie inside the ground truth of the manually labeled brains. Brain coverage is measured by the percentage of the

Table 1. The accuracy of the proposed localization approach at different dictionary sizes $k = 50$, $k = 100$, $k = 400$ and $k = 800$.

	k = 50	k = 100	k = 400	k = 800
Brain detection (% subjects)	87.27	89.09	94.55	90.91
Center in brain (% subjects)	80	81.82	98.18	98.18
Brain coverage ($\mu \pm \sigma\%$)	85.59 \pm 30.51	85.9 \pm 28.2	90.03 \pm 16.63	90.54 \pm 15.17
Dice coefficient ($\mu \pm \sigma\%$)	61.07 \pm 26.94	63.22 \pm 25.47	71.96 \pm 19	73.62 \pm 15.9
Average train time (minutes)	31.55	50.25	256.09	424.93
Average test time (minutes)	4.63	4.85	6.31	8.36

manually-labeled brain that are covered by the segmented boundary box. We have also used the dice coefficient [6] in order to measure the segmentation accuracy of the proposed approach. Table 1 shows the accuracy of our localization approach at different dictionary sizes $k = 50$, $k = 100$, $k = 400$ and $k = 800$. These results shows that increasing the dictionary size or the histogram bins (sparsity) increases the dice accuracy. However, it also increases significantly the processing time of training and testing. These experiments were done using parallel processing on a CPU with 32-cores and 128 GB RAM (Fig. 5).

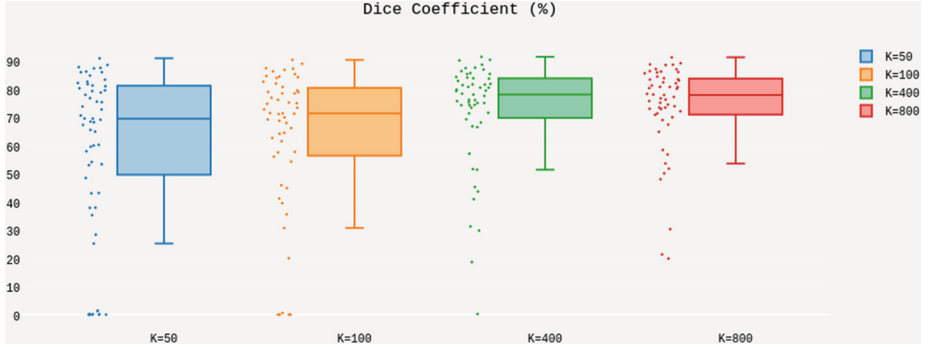


Fig. 5. The accuracy of the proposed localization approach at different dictionary sizes $k = 50$, $k = 100$, $k = 400$ and $k = 800$.

Our brain localization approach have achieved a 94.55% detection accuracy. It also could detect the center of the brain with prediction accuracy 98.18% of the test subjects while in [9] they achieved only 81%, 78% and 60% using coronal, axial and sagittal training data. In addition, the proposed approach does not depend on the orientation of the acquired data and it does not use any previous landmarks as in [9]. [12] achieved 100% detection accuracy of the brain; their method, however, requires previous information about the gestational age of the fetus to find the expected size of the brain. This information is later used to remove the outliers of the detected brain mask. Our proposed method has

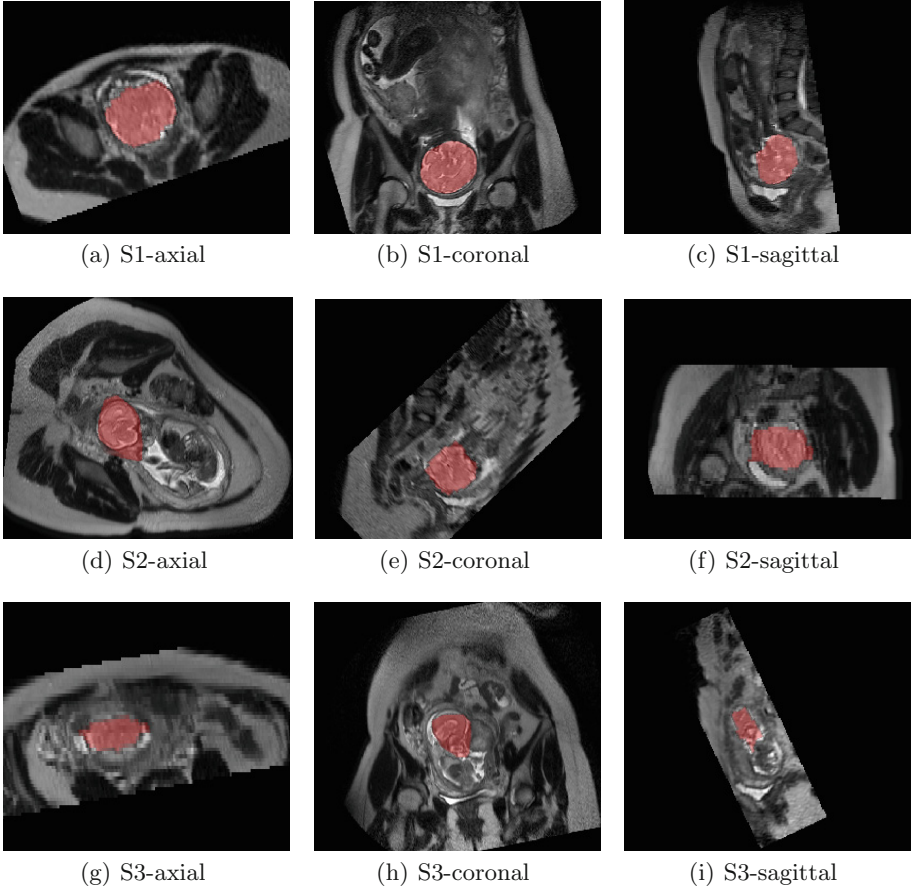


Fig. 6. The segmented brain at different cross sections (axial, coronal, sagittal) for three different test subjects S1, S2 and S3. The dice coefficient of S1=91.22 %, S2=70.01 %, and S3=56.79 %.

an advantage to be generic and does not require any prior information. Figure 6 shows the segmented brain at different cross sections for three test subjects with different dice accuracy.

4 Discussion and Conclusion

We have developed an automatic framework for localizing the brain in fetal MRI scans using superpixel graphical models. Superpixels have enabled the proposed detection algorithm to be faster and more efficient than using pixels for classification. Also, extending the extracted features from the individual superpixels to include features from the neighbors using superpixel graphical models, have provided more information about the image context instead of using only the

local information. The evaluation results achieved 94.55 % accuracy for the brain detection, which shows the potential of extending the proposed approach using superpixel graphs to segment other fetal organs such as the heart, lung, and placenta. According to the recent studies [5], the placental functions affect the birth weight as the placenta controls the nutrients transmissions from the maternal to the fetal circulation. Moreover, the extracted brain can be used for developing automatic motion correction and registration techniques for fetal MRI.

Acknowledgments. Thanks for the volunteer subjects and radiographers from St. Thomas Hospital London for the image acquisitions. We used the Medical Imaging Interaction Toolkit (MITK) [17] to visualize some of the figures. Amir Alansary is supported by the Imperial College PhD Scholarship.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2274–2282 (2012)
2. Anquez, J., Angelini, E.D., Bloch, I.: Automatic segmentation of head structures on fetal MRI. In: *Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 109–112. IEEE Press (2009)
3. Barber, C.B., Dobkin, D.P., Huhdanpaa, H.: The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw. (TOMS)* **22**(4), 469–483 (1996)
4. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
5. Damodaram, M., Story, L., Eixarch, E., Patel, A., McGuinness, A., Allsop, J., Wyatt-Ashmead, J., Kumar, S., Rutherford, M.: Placental MRI in intrauterine fetal growth restriction. *Placenta* **31**(6), 491–498 (2010)
6. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* **26**(3), 297–302 (1945)
7. Fulkerson, B., Vedaldi, A., Soatto, S.: Class segmentation and object localization with superpixel neighborhoods. In: *2009 IEEE 12th International Conference on Computer Vision*, pp. 670–677. IEEE (2009)
8. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L.: *Multivariate Data Analysis*. Pearson, Prentice Hall, Upper Saddle River (2006)
9. Ison, M., Donner, R., Dittrich, E., Kasprian, G., Prayer, D., Langs, G.: Fully automated brain extraction and orientation in raw fetal MRI. In: *Proceedings of the Workshop on Paediatric and Perinatal Imaging, MICCAI*, vol. 12, pp. 17–24 (2012)
10. Keraudren, K., Kuklisova-Murgasova, M., Kyriakopoulou, V., Malamateniou, C., Rutherford, M., Kainz, B., Hajnal, J., Rueckert, D.: Automated fetal brain segmentation from 2D MRI slices for motion correction. *Neuroimage* **101**, 633–643 (2014)
11. Keraudren, K., Kainz, B., Oktay, O., Kyriakopoulou, V., Rutherford, M., Hajnal, J.V., Rueckert, D.: Automated localization of fetal organs in MRI using random forests with steerable features. In: *Medical Imaging Computing and Computer-Assisted Intervention–MICCAI (2015)* (in press)

12. Keraudren, K., Kyriakopoulou, V., Rutherford, M., Hajnal, J.V., Rueckert, D.: Localisation of the brain in fetal MRI using bundled SIFT features. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013, Part I. LNCS, vol. 8149, pp. 582–589. Springer, Heidelberg (2013)
13. Lowe, D.G.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, vol. 2, pp. 1150–1157. IEEE (1999)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
15. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(10), 1744–1757 (2010)
16. Vedaldi, A., Fulkerson, B.: VLFeat: an open and portable library of computer vision algorithms. In: Proceedings of the International Conference on Multimedia, pp. 1469–1472. ACM (2010)
17. Wolf, I., Vetter, M., Wegner, I., Böttger, T., Nolden, M., Schöbinger, M., Hasenteufel, M., Kunert, T., Meinzer, H.P.: The medical imaging interaction toolkit. *Med. Image Anal.* **9**(6), 594–604 (2005)

Machine Learning Meets Medical Imaging

First International Workshop, MLMMI 2015, Held in

Conjunction with ICML 2015, Lille, France, July 11, 2015,

Revised Selected Papers

Bhatia, K.K.; Lombaert, H. (Eds.)

2015, X, 105 p. 31 illus. in color., Softcover

ISBN: 978-3-319-27928-2