Automated placenta segmentation from 3D ultrasound images

Ipek Oguz[1], Alison M. Pouch[1], Natalie Yushkevich[1], Hongzhi Wang[2], James C. Gee[1], Nadav Schwartz[3], and Paul A. Yushkevich[1]

[1] Penn Image Computing and Science Laboratory (PICSL), Department of Radiology, University of Pennsylvania, Philadelphia, PA
[2] IBM Almaden Research Center, Almaden, CA
[3] Maternal and Child Health Research Program, Department of OBGYN, University of Pennsylvania, Philadelphia, PA

Abstract. Quantifying placental volume and morphology is important for the study of adverse pregnancy outcomes and fetal programming. However, robust segmentation of the placenta in 3D ultrasound images remains a challenging task because of the variability of the shape, position and appearance of the placenta in these images, as well as the high level of noise. Current approaches require substantial user interaction to initialize segmentation. We propose the first fully automated approach to segment the placenta from 3D US images, using the multi-atlas label fusion framework. Multi-atlas label fusion methods have gained popularity in recent years given their high performance in a variety of challenging segmentation tasks. To obtain accurate registration between unseen target images and manually annotated atlases for placenta segmentation, we propose an initialization scheme that automatically aligns the atlases and target image using 3DUS cone alignment and image denoising. Then, a bootstrapping approach is implemented to boost the segmentation accuracy by using the results of the first round of multi-atlas label fusion to improve initialization of a second round of multi-atlas label fusion. We evaluate our method in a dataset of 13 subjects in first trimester of pregnancy with anterior placentas, in a leave-one-out study. The Dice overlap between our proposed algorithm and expert manual segmentations was $83.2 \pm 5.3\%$, on par with existing interactive approaches.

Keywords: segmentation, multi-atlas label fusion, placenta, ultrasound

1 Introduction

Adverse pregnancy outcomes, including preeclampsia and intrauterine growth restriction, are significant contributors to perinatal morbidity and mortality, and are believed to have their roots in abnormal placental development. Clinicians have observed that placental size and shape are significantly associated with perinatal outcomes and fetal programming [3, 4, 2]. In particular, 2D and 3D ultrasound has been used to study early, in utero placental size and morphology in relation to birthweight and preeclampsia [16, 15]. As such, placental morphology is crucially important for assessing gross placental development in vivo.

Nevertheless, quantification of the placenta from 3D ultrasound images is a challenging task, especially in early pregnancy when the appearance of the placenta and uterine tissue are not well differentiated. This lack of contrast makes it difficult to accurately detect the boundary, especially given the noisy nature of 3DUS images. Additionally, the shape and relative location of the placenta are highly variable, which in some cases make it difficult to even detect the location of the placenta automatically, much less determine its precise boundaries. Uterine contractions can also dramatically affect the shape of the placenta.

As a result of these challenges, the current techniques for quantifying placental volume and morphology from 3D ultrasound images require substantial amount of manual input from the user, which can be cumbersome and poorly standardized. These include the commercial Virtual Organ Computer-aided AnaLysis (VOCAL, GE Healthcare) software as well as a recent random-walker algorithm [6, 17]. In contrast, we propose a fully automated segmentation technique to maximize reliability and facilitate the integration of this technique into bedside clinical care. Our proposed technique leverages the recently popular multi-atlas segmentation technique [8, 19], which fuses information from a library of expert-labeled example datasets and finds image boundaries where there is little or no intensity contrast between the structure of interest and the surrounding structures. This approach has been applied successfully to many such challenging segmentation problems such as identifying hippocampal subfields in MRI of the brain [20, 18] and the mitral valve in pre-operative 3D echocardiography [12] and is particularly relevant for placental imaging where maternal habitus and fetal shadowing artifact can obscure these boundaries. While the placenta shape and appearance are highly variable, we hypothesized that the multi-atlas label fusion would nevertheless be able to produce accurate segmentations given a large enough atlas set and a careful registration scheme.

2 Methods

The main contribution of this manuscript is the application of the multi-atlas label fusion approach to the task of segmenting the placenta from US images. While the individual components of the algorithm (such as deformable registration or label fusion) are not novel, our approach is nevertheless highly innovative as it is the first fully automated placenta segmentation technique. An additional advantage of our approach is that it relies exclusively on open-source software that is readily available to the placental imaging community.

2.1 Multi-atlas label fusion framework

Registration-based segmentation approaches have long been popular strategies in analyzing anatomical structures in medical images. Given an expert-annotated atlas image, a target image can be segmented by deformably registering the atlas to the image and applying the same deformation to the annotation. However, for challenging segmentation tasks, there may be a large variability in the shape and appearance of the structure being segmented, decreasing the generalizability of a single atlas; alternatively, there may be weak boundaries between neighboring

structures either due to lack of contrast or image noise, which may make an accurate registration difficult to achieve. Multi-atlas label fusion (MALF) methods have recently been gaining popularity [8] as a general-purpose framework for such challenging segmentation problems. As the name suggests, the general idea is to leverage not one but multiple annotated atlases to segment an unseen target image. Each atlas is deformably registered to the target image, and the annotation for the atlas is deformed accordingly to generate a candidate segmentation. The label fusion step seeks to combine these alternative solutions, typically via a weighted voting scheme, into a single consensus segmentation. A variety of techniques have been proposed to assign weights to each candidate, ranging from uniform weights for a simple majority vote approach to more sophisticated schemes that aim to estimate the accuracy of each candidate by taking into account the similarity between the atlas and the target image. Even when some registrations fail or when some atlases are very different than the target image, these techniques can still often achieve a good segmentation by minimizing the weight of these poorly-matched candidates and instead focusing on the well-matched candidates.

For the placenta segmentation task, we propose to use the joint label fusion algorithm [20, 19]. This algorithm takes into account not only the similarity of each atlas to the target image, but also the similarity between each pair of atlases. The rationale for this approach is that pairs of atlases that are similar to each other are likely to generate similar candidate segmentations for the target image and can be seen as redundant to some extent. If these are poorly matched candidates, this can reduce the accuracy of the label fusion result. In contrast, the joint label fusion approach jointly estimates the weights for the atlases, which minimizes the redundancy from correlated atlases and leads to a more robust segmentation result. This approach is particularly well-suited to the placenta segmentation problem given the large variability of the placenta shape and appearance; some target images will only have few well-suited atlases with similar characteristics, and minimizing the impact from the remaining atlases is imperative for satisfactory segmentation results. Fig. 1 illustrates this process.

Our overall approach consists of 3 main stages, as illustrated in Fig. 2. We begin by rotating each image such that the axial direction of the ultrasound beam is aligned with the z axis of the 3D image and applying a median filter to reduce noise, in order to facilitate the subsequent linear and deformable registration steps. Then, each atlas image is registered to the target image and we apply the joint label fusion algorithm to generate an estimate for the placenta region in the target image. We repeat the entire process by using this estimate as a region of interest to maximize the registration accuracy in and around the placenta. Each label fusion stage is followed by a post-processing step to remove any geometric artifacts such as holes or disconnected pieces.

2.2 Pre-processing: US cone alignment and denoising

We begin by roughly aligning the image cones for each subject to facilitate the subsequent registration steps. We threshold the image at 0 to obtain the ultrasound cone and extract the surface of this region. The surface normals of

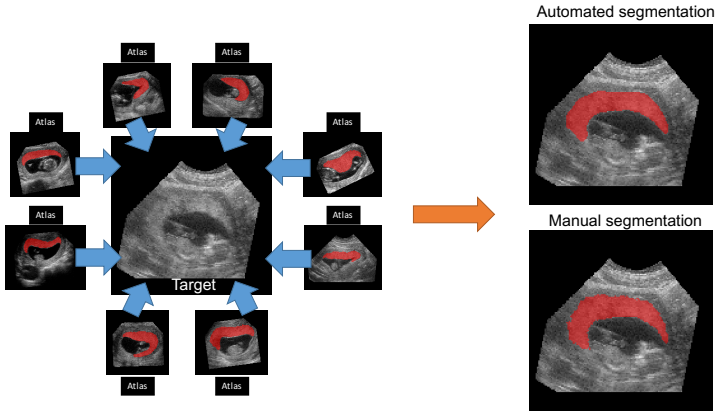


Fig. 1. The joint label fusion process illustrated on the subject with the median Dice score (84.9). Each atlas image is mapped (blue arrows) to the target image to generate a candidate segmentation. Given the large natural variability in the placenta shapes, no single atlas will map perfectly to each target image. However, our multi-atlas approach allows us to capture this variability and produces robust segmentation results.

this cone are computed, and the average normal direction is the main axis of the cone. We rotate the image such that this direction is aligned with the z axis of the image, and we translate it such that the center of mass of the cone is located at the origin. We apply a median filter with a neighborhood radius of 2 voxels to reduce the amount of noise in the image. While this may lead to the loss of fine-level details in the segmentation results, we found that this is well-suited for this application because the placenta has a very smooth surface.

2.3 ROI estimation

Given a pre-processed target image and a set of pre-processed atlases, we register each atlas to the target image. To this end, we perform a rigid registration followed by a deformable registration. Given the high variability of the anatomy and the noisy nature of the images, affine registration was found to be somewhat unstable with respect to scale parameters and therefore skipped in this stage. Normalized cross correlation was used as the similarity metric for all registration steps. Registration used an open-source implementation of greedy diffeomorphic image registration [1, 9] with fast cross-correlation metric computation. The manual segmentations of the atlases were then deformed into the target image space using the associated deformation fields, and the joint label fusion algorithm [20] was used to combine these into a single consensus segmentation. Mathematical morphology operators were used to remove topological defects. Specifically, holes in the object segmentation were detected and filled, and erroneous islands and protrusions were removed by eroding the segmentation by a 1-voxel radius, extracting the largest connected component, and dilating back up with the same structuring element. Finally, we dilate the segmentation result

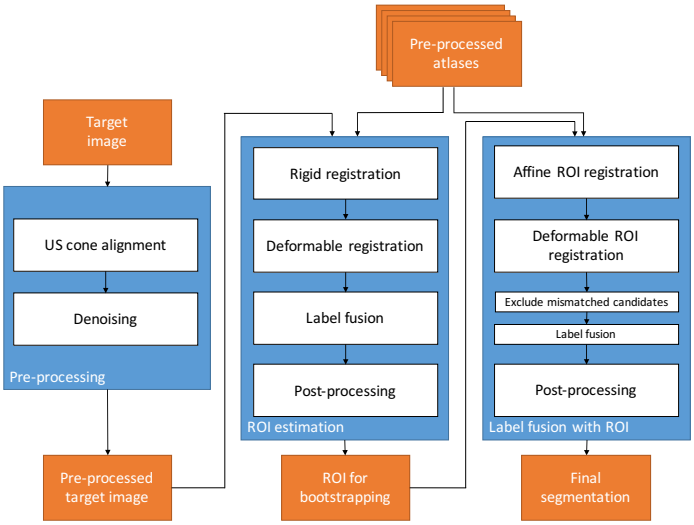


Fig. 2. The placenta segmentation workflow. We begin by rigidly aligning each image such that the US cone is aligned with the z axis and apply a median filter to reduce noise, in order to facilitate the subsequent linear and deformable registration steps. Then, each atlas image is registered to the target image and we apply the joint label fusion algorithm to generate an estimate for the placenta region in the target image. We repeat the entire process by using this estimate as an ROI to maximize the registration accuracy in and around the placenta. Finally we apply a post-processing step to remove any geometric artifacts such as holes or disconnected pieces.

by 10 voxels in each direction, in order to generate an ROI that includes any under-segmented portions of the placenta.

2.4 Final segmentation

For the final segmentation, we perform the same general steps as the ROI estimation, but in this stage we use the estimated ROI as a mask for registration. The rationale for this repetition is that registration algorithms typically perform better when they can be focused to a more 'standardized' problem such that the moving and fixed images contain roughly the same anatomy. For example, skull-stripped brain images tend to be easier to register accurately, compared to raw images that may contain various amounts and types of non-brain tissue.

The target image is masked using the dilated ROI described above, and each atlas is masked using the corresponding manual segmentation. We perform an affine registration between these two masked images using the normalized cross-correlation metric in a large neighborhood window ($10 \times 10 \times 10$ voxels), with a mild smoothing during the greedy update step (1.732 voxels by 0.7071 voxels). The deformable registration is performed on the unmasked images, but instead the dilated ROI is used to mask the gradient computation during the optimization. The image similarity term used was the normalized cross-correlation with

a smaller neighborhood ($4 \times 4 \times 4$ voxels), and more smoothing was applied during the greedy update step (5 voxels by 2 voxels). Any failed registrations, e.g. caused by a strong dissimilarity of the placentas between the atlas and the target, were detected and removed. To this end, the overlap between the ROI estimated in the previous stage and each candidate segmentation generated in this stage was computed, and the candidates in poor agreement with the ROI (Dice < 0.4) were excluded. Then, the joint label fusion algorithm was used to combine the remaining candidate segmentations. Holes in the object segmentation were detected and filled to generate the final segmentation.

3 Experimental Methods

3.1 Dataset

For our study, 14 women in first trimester of pregnancy were imaged with GE Voluson E8 ultrasound machines. Each image had isotropic resolution (mean: 0.47mm, min: 0.34mm, max: 0.59mm). The images were exported in DICOM format and converted to NIFTI during anonymization. Only subjects with anterior placentas were chosen for this study. One subject was excluded since the placental position differed from the others by more than 90 degrees.

3.2 Expert manual segmentation

Each of the 14 subjects were manually segmented by Anon. under the supervision of Anon., who has over 10 years of experience in prenatal ultrasound imaging and who has segmented 100's of placentas for other research endeavors. Anon. provided guidance for Anon.'s training and performed final inspection and approval for each segmentation. The publicly available ITK-SNAP software[4] [22] was used for segmentation. In each case, the perimeter of the placenta was manually traced in multiple slices in all 3 planes. The border of the fetal side of the placenta was easily visualized as there is a stark contrast between the anechoic amniotic fluid within the amniotic cavity and the echogenic placental mass. The border on the maternal side of the placenta was less pronounced, although the slightly more echogenic placental tissue was generally discernible from the surrounding myometrium. While the sparse segmentation was only performed approximately every 5 slices, the adjacent slices as well as the other 2 planes were able to inform the manual segmentation when the borders were less clear. The sparse segmentation was inspected by Anon. and any corrections made prior to a final inspection and approval. This manually defined sparse manual segmentation was then interpolated to create a smooth label map of the entire placenta, which was also reviewed by the manual experts.

3.3 Evaluation criteria

We evaluate the performance of our proposed algorithm in a leave-one-out validation study. For each of the 13 images, we use the remaining 12 images as

[4] `www.itksnap.org`

	Dice Overlap	Jaccard Overlap	Volume Similarity	False Positive Rate	False Negative Rate
ROI estimate with JLF	0.77 ± 0.07	0.63 ± 0.09	-0.03 ± 0.22	0.21 ± 0.09	0.23 ± 0.12
ROI estimate with MV	0.66 ± 0.11	0.51 ± 0.12	-0.25 ± 0.33	0.23 ± 0.09	0.39 ± 0.17
Final segmentation	0.83 ± 0.05	0.72 ± 0.08	0.01 ± 0.19	0.16 ± 0.10	0.16 ± 0.09

Table 1. Volumetric accuracy measures for the ROI estimate (pre-dilation) and final segmentation stages compared to the manual expert segmentations. Average and standard deviation of each measure is reported.

	Auto to Manual	Manual to Auto
ROI estimate	2.88 ± 1.46	2.59 ± 1.27
Final segmentation	1.95 ± 1.06	1.75 ± 0.88

Table 2. Surface-to-surface error measures between the automated and manual expert segmentations. The ROI esimate is pre-dilation. The average and standard deviation of distance is reported in mm, from automated to manual and vice versa.

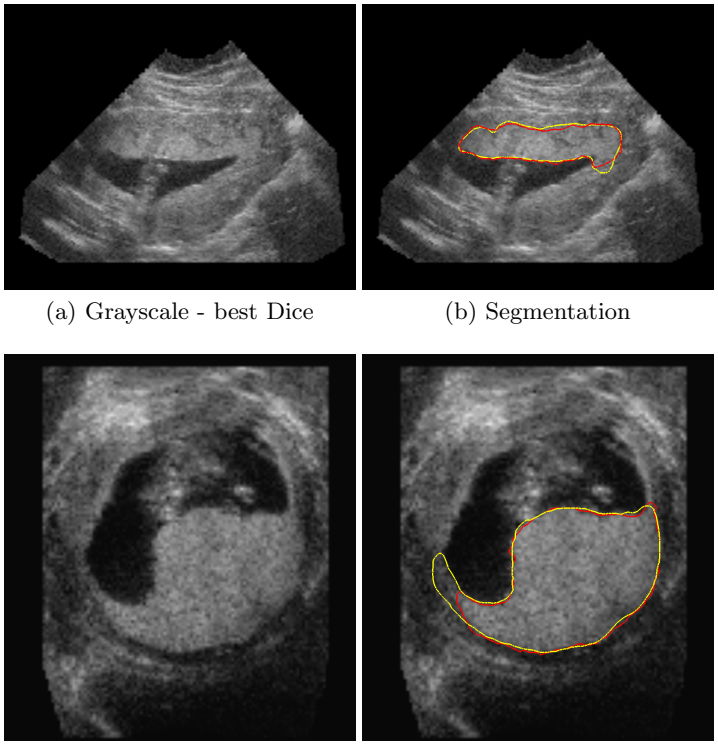
atlases to segment the target, and compare the automated segmentation to the expert manual segmentation. Specifically, we report the Dice and Jaccard overlap measures as well as the false positive and false negative rates. We also report the overall volume similarity ($2\frac{V_{auto}-V_{manual}}{V_{auto}+V_{manual}}$), as well as surface-to-surface distances. To compute the surface-based error measures, we used the marching cubes algorithm without smoothing to generate triangle mesh representations of the placenta surfaces. Additionally, to illustrate the impact of the joint label fusion method, we compare the ROI estimation with JLF to ROI estimation with a simple majority voting scheme.

4 Results

Tables 1 and 2 summarize the volumetric and surface-based accuracy measures, respectively. The overall accuracy of our fully automated segmentation approach was found to be on par with the interactive approaches described in [6, 17], where a Dice score of 0.86 was reported. We note that the third stage of our approach, i.e. the bootstrapping of the registration scheme using the estimated ROI's, offers a substantial boost in accuracy. The joint label fusion also offers a substantial improvement over a simplistic majority voting scheme. Fig. 3 presents qualitative results for the subjects with best and worst Dice scores.

5 Discussion and Conclusion

We presented preliminary results for the first fully automated approach to the segmentation of placentas from 3D ultrasound images. Compared to interactive methods, automated methods offer excellent reproducibility and may facilitate the integration into bedside clinical care. The availability of an automated tool is thus expected to enhance our understanding of early placental development



(a) Grayscale - best Dice (b) Segmentation

(c) Grayscale - worst Dice (d) Segmentation

Fig. 3. Qualitative segmentation results for the subject with the best (a-b) and worst (c-d) Dice score in the dataset. Red: manual, yellow: automated.

and aid the study of environmental factors, such as maternal nutrition, that may affect placental development and fetal growth.

Of the existing semi-automated placenta segmentation approaches, the commercial VOCAL software requires 6-12 successive manual tracings of the placental perimeter, limiting the potential for placental biometry to be incorporated as a bedside clinical tool. Moreover, interpolation of 3D placental shape from only 6-12 2D slices is a likely source of error in placenta volume assessment with this tool. Despite such limitations, studies have used VOCAL to demonstrate a strong, independent association between placental measures at 11-14 weeks with subsequent delivery of a small-for-gestational-age [16, 15] or macrosomic infant [14] supporting the notion that placental size, even as early as the first trimester, is a significant indicator of the functional potential of the placenta. However, as with most if not all interactive image segmentation approaches, the reproducibility of the VOCAL approach is problematic; for example, the reported mean placental volume calculated by VOCAL at 11-14 weeks ranges anywhere from 50 - 75cc [16, 15, 7, 11, 13, 21, 10].

Recently, another semi-automated technique for placenta segmentation, using a random-walker algorithm, from a 3DUS volume set has been developed [6, 17]. However, this approach requires the user to follow a specific protocol for placing 'seed' regions to define the placenta-uterine boundary, followed by additional manual input to define the placental surface [5], and it can be quite sensitive to deviations from this protocol [17]. The reliance on manual user input may lead to similar issues and biases as VOCAL.

Thus, a standardized, reproducible, automated and validated tool to quantitatively assess 3D placental morphology at the bedside is needed.

We evaluated the proposed fully automated method using a leave-one-out approach in a dataset of 13 patients. Based on our segmentation assessment (Dice overlap = 83 ± 5), the performance of our approach is on par with the only non-commercial semi-automated strategy presented in [17], which reports Dice overlap of 86 ± 6, albeit on a larger data set. The absolute volume similarity achieved by our method was -1.32 mL on average, compared to -0.78mL and -6.2mL for the Random Walker (for two observers) and -1.53mL and -9.59mL for VOCAL, reported in [17]. Given the importance of the overall placental volume as a morphological endpoint in a variety of studies, this result is especially promising.

We note that since the interactive random walker algorithm relies on the user input for determining the location and appearance of the placenta, it is readily applicable to both anterior and posterior placentas, whereas our approach is currently limited to anterior placentas only. Extension of the proposed approach to posterior or lateral placentas will likely require a more robust rigid/affine registration scheme to detect the correct location of the placenta and/or an automated atlas selection technique to identify atlases with similarly positioned placentas. This extension as well as further evaluation on images from women in their second and third trimester of pregnancy remain as future work.

Acknowledgments. This work was funded by the Human Placenta Project (U01 HD087180), as well as NIH grants R01 EB017255 and F32 HL119010.

References

1. Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Medical Image Analysis 12(1), 26–41 (2008)
2. Baptiste-Roberts, K., Salafia, C.M., Nicholson, W.K., Duggan, A., Wang, N.Y., Brancati, F.L.: Gross placental measures and childhood growth. The Journal of Maternal-Fetal & Neonatal Medicine 22(1), 13–23 (2009)
3. Barker, D.J., Bull, A.R., Osmond, C., Simmonds, S.J.: Fetal and placental size and risk of hypertension in adult life. BMJ 301(6746), 259–262 (1990)
4. Biswas, S., Ghosh, S.K.: Gross morphological changes of placentas associated with intrauterine growth restriction of fetuses: A case control study. Early Human Development 84(6), 357–362 (2008)
5. Collins, S.L., Stevenson, G.N., Noble, J.A., Impey, L.: Elsevier Trophoblast Research Award Lecture: Searching for an early pregnancy 3-D morphometric ultrasound marker to predict fetal growth restriction. Placenta 34, S85–S89 (2013)

6. Collins, S.L., Stevenson, G.N., Noble, J.A., Impey, L.: Rapid Calculation of Standardized Placental Volume at 11 to 13 Weeks and the Prediction of Small for Gestational Age Babies. Ultrasound in Medicine & Biology 39(2), 253–260 (2013)
7. Hafner, E., Metzenbauer, M., Höfinger, D., Munkel, M., Gassner, R., Schuchter, K., Dillinger-Paller, B., Philipp, K.: Placental Growth from the First to the Second Trimester of Pregnancy in SGA-foetuses and Pre-eclamptic Pregnancies Compared to Normal Foetuses. Placenta 24(4), 336–342 (2003)
8. Iglesias, J.E., Sabuncu, M.R.: Multi-atlas segmentation of biomedical images: A survey. Medical Image Analysis 24(1), 205–219 (2015)
9. Joshi, S., Davis, B., Jomier, M., Gerig, G.: Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 23 Suppl 1, S151–60 (2004)
10. Law, L.W., Leung, T.Y., Sahota, D.S., Chan, L.W., Fung, T.Y., Lau, T.K.: Which ultrasound or biochemical markers are independent predictors of small-for-gestational age? Ultrasound in Obstetrics & Gynecology 34(3), 283–287 (2009)
11. Odibo, A.O., Goetzinger, K.R., Huster, K.M., Christiansen, J.K., Odibo, L., Tuuli, M.G.: Placental volume and vascular flow assessed by 3D power Doppler and adverse pregnancy outcomes. Placenta 32(3), 230–234 (2011)
12. Pouch, A.M., Wang, H., Takabe, M., Jackson, B.M., Sehgal, C.M., Gorman III, J.H., Gorman, R.C., Yushkevich, P.A.: Automated Segmentation and Geometrical Modeling of the Tricuspid Aortic Valve in 3D Echocardiographic Images. In: MICCAI, pp. 485–492 (2013)
13. Rizzo, G., Capponi, A., Cavicchioni, O., Vendola, M., Arduini, D.: Placental vascularization measured by 3D power Doppler ultrasound at 11 to 13 + 6 weeks gestation in normal and aneuploid fetuses. Ultrasound in Obstetrics & Gynecology 30(3), 259–262 (2007)
14. Schwartz, N., Quant, H.S., Sammel, M.D., Parry, S.: Macrosomia has its roots in early placental development. Placenta 35(9), 684–690 (2014)
15. Schwartz, N., Coletta, J., Srinivas, S., Pessel, C., Timor, I.E., Parry, S., Salafia, C.: Placental volume measurements early in pregnancy predict adverse perinatal outcomes. American Journal of Obstetrics and Gynecology 201(6), S142–S143 (2009)
16. Schwartz, N., Sammel, M.D., Leite, R., Parry, S.: First-trimester placental ultrasound and maternal serum markers as predictors of small-for-gestational-age infants. American Journal of Obstetrics and Gynecology 211(3), 253.e1–253.e8 (2014)
17. Stevenson, G.N., Collins, S.L., Ding, J., Impey, L., Noble, J.A.: 3-D Ultrasound Segmentation of the Placenta Using the Random Walker Algorithm: Reliability and Agreement. Ultrasound in Medicine & Biology 41(12), 3182–3193 (2015)
18. Wang, H., Das, S.R., Suh, J.W., Altinay, M., Pluta, J., Craige, C., Avants, B., Yushkevich, P.A.: A learning-based wrapper method to correct systematic errors in automatic image segmentation: Consistently improved performance in hippocampus, cortex and brain segmentation. NeuroImage 55(3), 968–985 (2011)
19. Wang, H., Suh, J.W., Das, S.R., Pluta, J.B., Craige, C., Yushkevich, P.A.: Multi-Atlas Segmentation with Joint Label Fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 35(3), 611–623 (2013)
20. Wang, H., Yushkevich, P.: Multi-atlas segmentation with joint label fusion and corrective learning—an open source implementation. Frontiers in Neuroinformatics 7 (2013)
21. Wegrzyn, P., Faro, C., Falcon, O., Peralta, C.F.A., Nicolaides, K.H.: Placental volume measured by 3D ultrasound at 11 to 13 + 6 weeks of gestation: relation to chromosomal defects. Ultrasound in Obstetrics & Gynecology 26(1), 28–32 (2005)
22. Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C., Gerig, G.: User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. NeuroImage 31(3), 1116–1128 (2006)