

Deep learning-based segmentation of the placenta and uterus on MR images

Maysam Shahedi^a, Catherine Y. Spong^b, James D. Dormer^a,
Quyen N. Do^c, Yin Xi^{c,d}, Matthew A. Lewis^c, Christina Herrera^b,
Ananth J. Madhuranthakam^{c,e}, Diane M. Twickler^b and
Baowei Fei^{a,d,e,*}

^aUniversity of Texas at Dallas, Department of Bioengineering, Dallas, Texas, United States

^bUniversity of Texas Southwestern Medical Center, Department of Obstetrics and Gynecology, Dallas, Texas, United States

^cUniversity of Texas Southwestern Medical Center, Department of Radiology, Dallas, Texas, United States

^dUniversity of Texas Southwestern Medical Center, Department of Clinical Science, Dallas, Texas, United States

^eUniversity of Texas Southwestern Medical Center, Advanced Imaging Research Center, Dallas, Texas, United States

Abstract

Purpose: Magnetic resonance imaging has been recently used to examine the abnormalities of the placenta during pregnancy. Segmentation of the placenta and uterine cavity allows quantitative measures and further analyses of the organs. The objective of this study is to develop a segmentation method with minimal user interaction.

Approach: We developed a fully convolutional neural network (CNN) for simultaneous segmentation of the uterine cavity and placenta in three dimensions (3D) while a minimal operator interaction was incorporated for training and testing of the network. The user interaction guided the network to localize the placenta more accurately. In the experiments, we trained two CNNs, one using 70 normal training cases and the other using 129 training cases including normal cases as well as cases with suspected placenta accreta spectrum (PAS). We evaluated the performance of the segmentation algorithms on two test sets: one with 20 normal cases and the other with 50 images from both normal women and women with suspected PAS.

Results: For the normal test data, the average Dice similarity coefficient (DSC) was 92% and 82% for the uterine cavity and placenta, respectively. For the combination of normal and abnormal cases, the DSC was 88% and 83% for the uterine cavity and placenta, respectively. The 3D segmentation algorithm estimated the volume of the normal and abnormal uterine cavity and placenta with average volume estimation errors of 4% and 9%, respectively.

Conclusions: The deep learning-based segmentation method provides a useful tool for volume estimation and analysis of the placenta and uterus cavity in human placental imaging.

© 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.8.5.054001](https://doi.org/10.1117/1.JMI.8.5.054001)]

Keywords: convolutional neural network; deep learning; image segmentation; placenta; uterus; magnetic resonance imaging.

Paper 21065R received Mar. 31, 2021; accepted for publication Sep. 2, 2021; published online Sep. 25, 2021.

1 Introduction

The placenta is a critical and complex organ that provides oxygen and nutrition to the growing fetus and removes waste from its blood. Fetal health strongly depends on the functionality of the

*Address all correspondence to Baowei Fei, bfei@utdallas.edu

placenta and therefore any abnormality of the placenta could be harmful to the fetus and mother.^{1,2} Placenta previa, placenta accreta spectrum (PAS), and uteroplacental insufficiency are the most frequent indications for imaging studies during pregnancy.^{3–6} Assessment of the placenta *in vivo* across gestation is critical to understand placental structure, function, and development and to identify strategies to optimize pregnancy outcome.⁷

Two-dimensional (2D) ultrasound imaging is the clinical standard imaging modality used for accessing placental health and diagnosis of the abnormalities. In placenta imaging, the location, shape, and volume of the placenta are important factors used in detecting abnormalities.^{8–10} In addition, the interface of the placenta at its uterine implantation is an important component of placental assessments. In 2D ultrasound imaging, physicians create three-dimensional (3D) models in their mind and subjectively determine volume, location, and features of the placenta, which is a challenging task. The placenta size is usually estimated based on the longest diameter of its shape detected by the physicians in 2D views.¹¹ 3D imaging with magnetic resonance imaging (MRI) could address some of these issues. Studies in the literature have shown that 3D segmentation of placenta in magnetic resonance (MR) images could be used as a part of a clinical procedure for monitoring conditions that result in pregnancy and birth complications such as PAS, fetal growth restriction, and suspicion of intrauterine fetal demise.^{12–15} However, manual segmentation of the placenta in MRI is time-consuming with high inter- and intra-observer variability.¹⁶

Although there have been efforts to develop computerized algorithms to segment the placenta in MRI,^{17,18} either the accuracy of the presented methods was relatively low (Dice coefficient¹⁹ was about 72%¹⁸) or the computational time was high (1 to 2 min/volume in average) and required multiple image volumes (sagittal and axial acquisitions).¹⁷ The size of the dataset (16 images) was another limitation of a previous study.¹⁷ There are a few algorithms for uterine segmentation in MRI^{20,21} in nonpregnant women. Namías et al.²⁰ used a local binary pattern-based texture feature extraction to segment the uterus. They tested their algorithm on nine patients and reported a Dice coefficient of 81%. Kurata et al.²¹ used a deep learning algorithm for uterine segmentation in MRI using U-Net architecture. They tested their method on 122 MR images acquired from 72 patients with benign or malignant tumors and 50 healthy cases. They reported Dice coefficients of 84% and 78% for patients and healthy cases, respectively. They also reported the average mean absolute distance between their results and manual segmentation. The average mean absolute distance was reported in pixel unit. Since the voxel size was not consistent across their image dataset, using a physical unit (e.g., millimeter) could give a better understanding about the surface misalignment.

Deep learning-based approaches demonstrated a strong capability for fast segmentation of medical images with high accuracy. At the 2020 SPIE Medical Imaging Conference, we presented an interactive deep learning-based segmentation algorithm for segmenting the uterine cavity and placenta in MRI of normal patients.²² We used minimal operator input to improve the segmentation performance. This study is a comprehensive extension of our preliminary work to achieve fast, accurate, and repeatable segmentation of both uterus and placenta in both normal women and women with suspicion of PAS. Figure 1(a) shows a general block diagram of our segmentation method. We adopted the U-Net architecture to present a 3D end-to-end fully

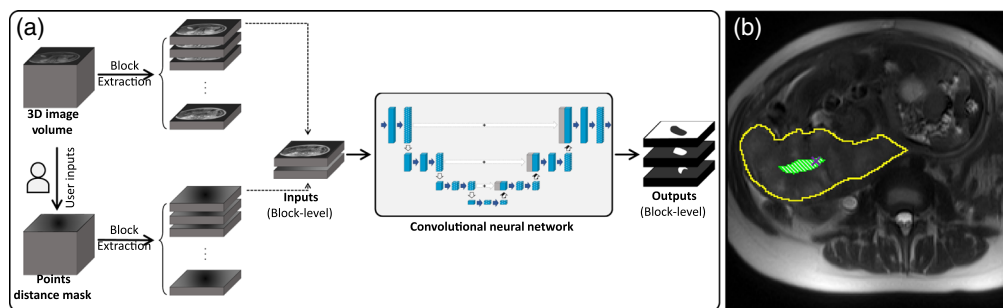


Fig. 1 (a) Schematic block diagram of the segmentation method. (b) Simulating user manual interaction. The yellow solid contour shows the placenta, the green, hashed area is the 5% placenta's central area, and the purple cross is the randomly selected point within that area.

convolutional neural network (CNN) for 3D image segmentation. We doubled the data size and tested the trained CNN model on the data that included images from subjects with both normal and abnormal placentas.

2 Methods

2.1 Data

In this study, we used two datasets of T2-weighted MR images of pregnant women throughout gestation (20 to 38 weeks). One dataset contained 100 3D uterus MRIs from 100 normal pregnant women. Each image volume had 28 to 52 2D transverse slices. For 99 cases, each slice was originally 256×256 pixels in size. For one case, the slice size was 212×256 pixels. We resized that image to 256×256 by zero-padding to make the slice size consistent across the dataset. The axial image slice spacing for all the images was 7.0 mm and the in-plane pixels were isotropic for all the slices with resolutions ranging from $1.0547 \times 1.0547 \text{ mm}^2$ to $1.7188 \times 1.7188 \text{ mm}^2$. The second dataset contained 99 3D uterus MRIs from 99 pregnant women. Each image volume had 40 to 62 2D transverse slices all 256×256 pixels in size. The axial image slice spacing for all the images was 7.0 mm and the pixels were isotropic for all the slices with resolutions ranging from $1.1719 \times 1.1719 \text{ mm}^2$ to $1.7578 \times 1.7578 \text{ mm}^2$. The subjects of this dataset were suspected of having PAS or other placenta abnormalities and 67 of them had postpartum hysterectomy operations.

Images of both datasets were acquired by 1.5 Tesla clinical MRI scanners (MAGNETOM Avanto and MAGNETOM Aera, Siemens, Erlangen, Germany) at the same medical center. A body array coil covered the entire pelvis for signal reception during imaging. For each image, the uterine cavity and placenta were manually segmented by an expert radiologist.

This study was approved by the research ethics board of our institution and written informed consent was obtained from all patients prior to enrolment.

2.2 User Interaction

Our previous observations suggested incorporating user interaction for higher segmentation accuracy, especially for the placenta segmentation, when a limited number of training images are available. Therefore, for each image, the operator selected the superior and the inferior axial slices of the uterine cavity. This information was used for image block extraction explained in the preprocessing subsection. In addition, the superior and the inferior axial slices of the placenta were selected for each patient, and the center of the placenta tissue was estimated on both slices. Then, three axial image slices, evenly spaced between the two selected slices for the placenta, were automatically selected and the operator was asked to define the approximate center of the placenta tissue on each of the three slices. Therefore, the operator needed to browse through the image slices and click seven times in total; twice for the uterus bounding slices and five times for the placenta bounding slices and the center points. In this study, we simulated the operator interaction by selecting a point defined for each slice randomly from the 5% central area of the axial cross-section of the placenta tissue on the slice [Fig. 1(b)]. Therefore, there are five points selected on placenta tissue. We use linear interpolation between the five selected points to define one point per placenta slice. To provide this information as an input, we made a binary image volume with the same size as the MR image and assigned ones to the voxels that are selected as center points and zeros to all the other voxels. We applied the 3D Euclidean distance transform to this binary image volume to make gradients toward the selected points. We used the result distance mask as the second input channel of the CNN.

2.3 Preprocessing

We applied a median filter to all the 2D image slices under a 3×3 window to reduce the noise while preserving the edges. Then, we normalized the voxel intensity distribution of each image by clipping the intensities on the 5th and 99.9th percentile followed by rescaling the intensities between zero and one, as shown in Eq. (1). This normalization of the intensity distribution helps

to reduce the effect of background voxels and makes the intensity distribution more consistent across the image dataset:

$$\hat{I}_i(x, y, z) = \begin{cases} 0, & I_i(x, y, z) < p_5(I_i) \\ 1, & I_i(x, y, z) > p_{99.9}(I_i) \\ \frac{I_i(x, y, z) - p_5(I_i)}{p_{99.9}(I_i) - p_5(I_i)}, & \text{otherwise} \end{cases} \quad (1)$$

where $I_i(x, y, z)$ represents the intensity of the i 'th image in the image dataset at coordinate (x, y, z) , and $p_5(I_i)$ and $p_{99.9}(I_i)$ are the 5th and 99.9th percentiles of the pixel intensities in the image, respectively. $\hat{I}_i(x, y, z)$ is the normalized image.

To reduce the intersubject variability in placenta intensity, we linearly scaled the intensities of the voxels to shift the average intensity (\bar{i}_i) of the selected center points and all the voxels around them with in-plane distances of less than three pixels to a constant intensity value (\bar{i}) across the image set [see Eq. (2)]. Here, the constant value is the average placenta intensity of the first image in the set:

$$\hat{I}_i(x, y, z) = \hat{I}_i(x, y, z) \times \frac{\bar{i}}{\bar{i}_i}, \quad (2)$$

where $\hat{I}_i(x, y, z)$ is the i 'th image after intensity scaling. In this study, $\bar{i} = 0.154$.

There is a full-overlap between the placenta and uterine cavity. To make the segmentation labels independent from each other with no overlap (one-hot encoding), we used the portion of the uterus that was not covered by the placenta as the uterine cavity channel. Then, we made a three-channel label that included background, uterine cavity (with no overlap with placenta), and placenta labels.

We used a block-based segmentation approach by extracting a set of 3D five-slice blocks from the images. Each block (B_i^k) contained five sequential 2D axial slices, with the uterus present in at least one of the slices. For each image, the adjacent blocks were defined with four-slice overlaps (Fig. 2). We applied the same process for extracting blocks from the second input channel.

2.4 Experiments

Experiment I: We randomly divided the images of the first dataset into 70 training, 10 validation, and 20 test images. We used data augmentation to double the number of the data using

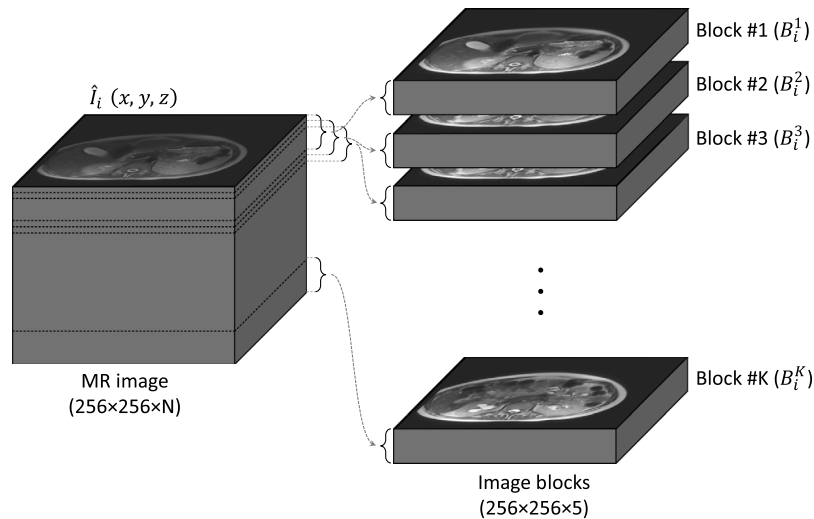


Fig. 2 Image blocks extraction of the 3D MR image.

left-to-right flipping of the images, yielding 140 training, 20 validation, and 40 test images. We used the data to train and test a CNN model (model I) for simultaneous segmentation of the uterus and placenta.

Experiment II: In this experiment we used all 199 images from both datasets to train and test the CNN (model II). We randomly chose 25 images from each dataset for testing (50 images in total) and used the remaining 149 images for training (129 images) and validation (20 images). This experiment tested the impact of adding images from abnormal placenta on the segmentation performance of the CNN model. Since in this experiment the number of training images was greater than experiment I, to avoid high computation load during training, we did not use any data augmentation for training and validation. We reported the average results for the whole test group and also for 35 nonhysterectomy cases and 15 hysterectomy cases separately.

2.5 Fully Convolutional Neural Network Architecture

In this study, we used a customized version of U-Net,²³ which is a fully convolutional neural network (FCNN). We modified the architecture to make it 3D and used that for multilabel image segmentation. Figure 3 shows the architecture of the proposed network. The network is a four-level U-Net model with 21 layers including 18 convolutional and three max pool layers. We kept the size of the output channels for all the convolutional layers the same as the input channels by zero-padding the input channel before convolution. For the two upper levels of the network, we used a convolution kernel size of $5 \times 5 \times 3$, and for the two lower levels we used a kernel size of $3 \times 3 \times 3$. We applied a dropout rate to the neurons of 10 layers shown in Fig. 3. The input has two channels: the 3D image blocks of the MRI and the points' distance mask. The output has three channels, one channel for uterine cavity, one for placenta, and one for background (i.e., anything not the uterus or placenta).

Due to an imbalance between the number of background and foreground (placenta and uterus) voxels, we used a loss function based on Dice similarity coefficient¹⁹ (DSC):

$$L = 1 - \frac{2 \sum_x \sum_y \sum_z [p(I_{x,y,z}) \cdot G_{x,y,z}]}{\sum_x \sum_y \sum_z [p(I_{x,y,z})] + \sum_x \sum_y \sum_z [G_{x,y,z}]}, \quad (3)$$

where $p(I_{x,y,z})$ is the probability value of the output probability map corresponds to the input image $(I_{x,y,z})$ at (x, y, z) and $G_{x,y,z}$ is the value of the reference binary mask at (x, y, z) . To avoid bias toward placenta or uterine cavity segmentation, we calculated two loss values, L_{uterus} for the uterine cavity and $L_{placenta}$ for placenta. We used the average of the losses as the total loss. The loss was calculated at the block-level during training. For optimization during training, we used the Adadelta²⁴ gradient-based optimizer.

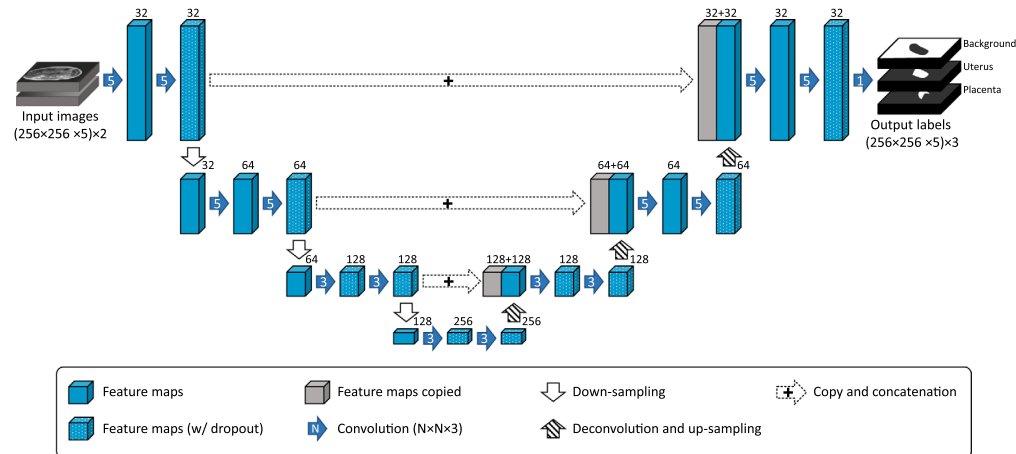


Fig. 3 The FCNN architecture (four-level 3D U-Net). The number of feature maps for each layer is mentioned at the top of the layer.

2.6 Implementation Details

We used the TensorFlow²⁵ machine learning system to implement the 3D U-Net model on a Python platform on a high-performance computer with 512 GB of memory and NVIDIA TITAN Xp GPU. We used a batch size of five (limited by the size of GPU memory) and an initial learning rate of 1.0, with a dropout rate of 40% (see Fig. 3), and the decay rate and epsilon conditioning parameters for Adadelta optimizer of 0.9 and 1×10^{-10} , respectively. The initial learning rate value was chosen based on a systematic search for an optimized learning rate around the default initial learning rate of Adadelta optimizer (i.e., 0.001). We trained the networks for up to 1000 epochs and chose the best trained model based on the highest validation accuracy.

2.7 Postprocessing and Evaluation

After testing the network on the 3D image blocks of an MR volume, we integrated the results by averaging the probabilities over the overlapped slices results to a probability map for the whole 3D image. We applied a similar process to the flipped version of the image to improve the segmentation result, using the average of the output probabilities corresponding to the original image and the flipped image as the final probability map. Thresholding was applied to build output binary masks out of the probability maps, which were then compared against manual segmentation labels. Our segmentation error metrics included the DSC, 95% Hausdorff distance (HD), and signed volume difference (ΔV):

$$\Delta V = V_{\text{seg}} - V_{\text{ref}}, \quad (4)$$

where V_{seg} is the volume of the object on the output segmentation label and V_{ref} is the volume of the object on the manual segmentation label. HD represents the largest distance between two surfaces within the 95% of the closest surface points. We reported DSC in percent, HD in mm, and ΔV in cm^3 and percent. Two-sample t -tests (two-tailed) were used to test the difference in mean DSC, HD, and ΔV between the first and second experiments. In each case, we tested the null hypothesis that the means of the metric resulting from the first experiment (experiment I) are equal to the means of the metric for the second experiment (experiment II). Significance level was set at 0.01.

3 Results

3.1 Training and Testing Results

We trained the first network (model I) until we reached the highest validation accuracy at epoch 530. The image block-level training and validation DSCs were 93.2% and 86.7%, respectively.

We segmented the test images using the trained CNN model. Table 1 shows the test results using the 20 test images of the first dataset. The metric values in the table are all at the full 3D image-level. The volume of the uterine cavity in the first test data ranged from 818 to 3993 cm^3 (mean \pm standard deviation [SD] = $2445 \pm 915 \text{ cm}^3$), whereas the placenta volume ranged from 161 to 1149 cm^3 (mean \pm SD = $629 \pm 268 \text{ cm}^3$) based on the provided manual segmentation labels. We also trained and tested the CNN on all the images from both datasets (model II). Table 2 shows the average test results across all 50 test cases as well as the hysterectomy and

Table 1 Test results of experiment I (average \pm SD).

Data	N	Uterine cavity				Placenta			
		DSC (%)	ΔV (cm^3)	ΔV (%)	HD (mm)	DSC (%)	ΔV (cm^3)	ΔV (%)	HD (mm)
Normal	20	92.0 ± 4.3	6 ± 254	2 ± 12	26.9 ± 14.1	81.9 ± 6.1	-53 ± 100	-9 ± 16	16.1 ± 7.4

Table 2 Test results of experiment II for the whole test set as well as hysterectomy (Hyst.) and nonhysterectomy (Non-hyst.) cases separately (average \pm SD). The values in bold show significant differences in comparison with test results in Table 1 ($p < 0.01$).

Data group	N	Uterine cavity				Placenta			
		DSC (%)	ΔV (cm ³)	ΔV (%)	HD (mm)	DSC (%)	ΔV (cm ³)	ΔV (%)	HD (mm)
Total	50	87.5 \pm 5.7	-247 \pm 589	-4 \pm 14	26.0 \pm 9.0	82.5 \pm 5.8	-12 \pm 151	0 \pm 15	25.2 \pm 22.6
Non-hyst.	35	88.6 \pm 4.9	-108 \pm 597	0 \pm 14	25.8 \pm 9.1	83.2 \pm 2.1	10 \pm 113	3 \pm 13	23.6 \pm 22.8
Hyst.	15	84.9 \pm 6.7	-571 \pm 434	-14 \pm 11	26.5 \pm 8.8	81.0 \pm 7.2	-62 \pm 210	-7 \pm 17	28.9 \pm 22.4

nonhysterectomy subsets, separately. The volume of the uterine cavity in the experiment II test set ranged from 818 to 7857 cm³ (mean \pm SD = 3165 \pm 1392 cm³), whereas the placenta volume ranged from 160.6 to 2239 cm³ (mean \pm SD = 929 \pm 425 cm³) based on the provided manual segmentation labels. In Table 2, the values in bold show where statistically significant differences were observed between the corresponding values in Tables 1 and 2 ($P < 0.01$). The total computational time for full 3D segmentation of one patient was about 25 s on average.

Figure 4 shows the qualitative results in 3D for four sample cases from experiment I. Figure 5 shows the segmentation results for the same four cases on 2D axial slices. Figures 6 and 7 show the qualitative results in 3D and 2D, respectively, for four sample cases from experiment II.

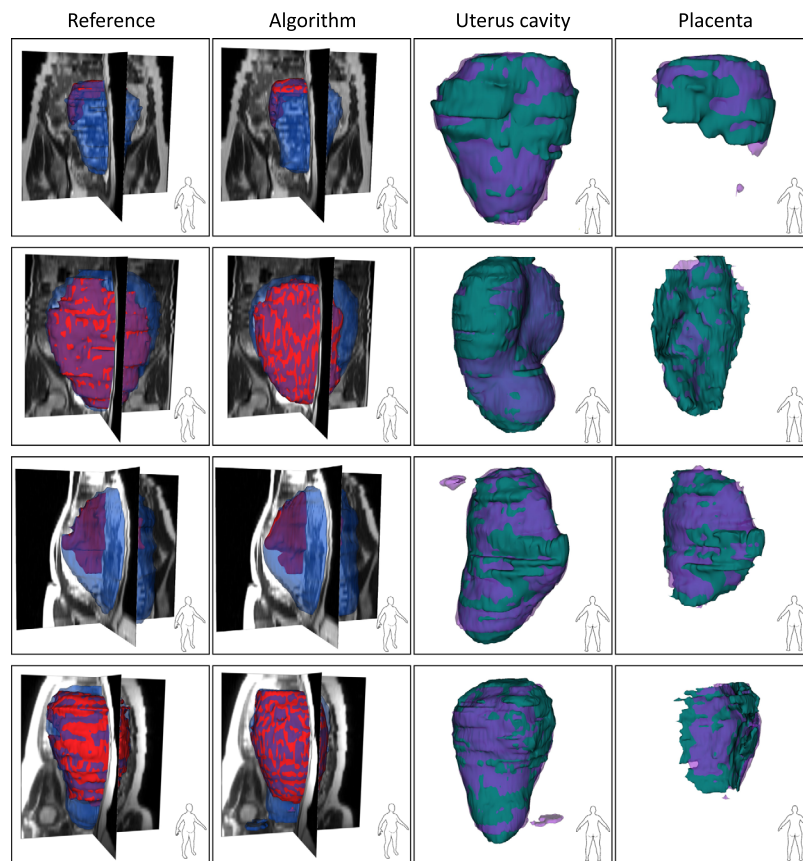


Fig. 4 Qualitative results of the uterus and placenta segmentation of four sample images from the first test set shown in 3D. Each row shows the results for one patient. For the first two columns, the blue, semitransparent shapes show the uterine cavity and the solid, red shapes show the placenta. For the last two columns, the green, solid shapes show the manual segmentations and the semi-transparent, purple shapes show the algorithm segmentation results.

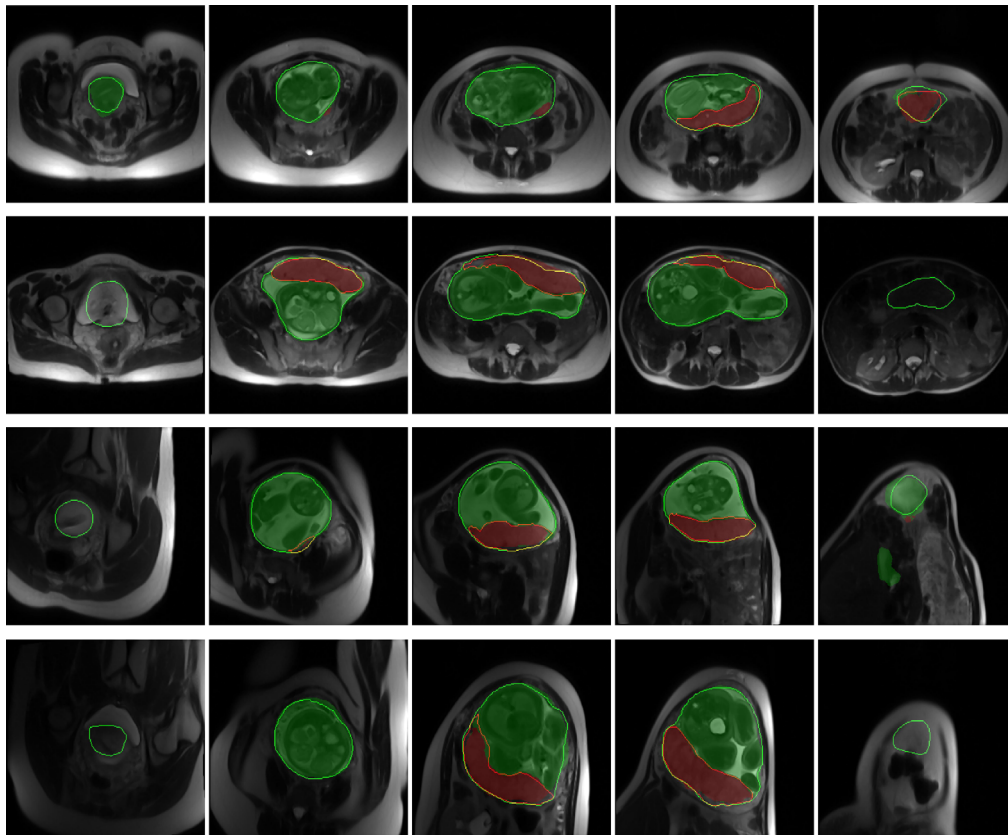


Fig. 5 Qualitative results of the uterus and placenta segmentation in 2D for the four subjects are shown in Fig. 4. Each row shows the results for one patient on five sample axial slices from (left) inferior to (right) superior. The semitransparent green (bright) and red (dark) regions show the algorithm segmentation results for the uterine cavity and placenta, respectively. The solid green (bright) and red (dark) contours show the manual segmentation provided by the expert observer for the uterine cavity and placenta, respectively.

During validation and test, we observed some small isolated regions in the segmentation labels of the uterine cavity and placenta (see the first, third, and fourth rows of Fig. 4) that looked like false positive components. As a postprocessing step, we automatically removed them by keeping only the largest 3D segmentation component. Adding the postprocessing step improved the average DSC values about 0.1% for both uterine cavity and placenta in experiment I. The postprocessing did not improve the uterine cavity segmentation in experiment II and improved the placenta segmentation by about 0.2% in terms of DSC.

4 Discussion

The proposed deep learning approach for multiclass 3D segmentation segments the uterine cavity and placenta in MR image volumes simultaneously with high segmentation accuracy. The high intersubject variability in placenta and uterus size, shape, and appearance is one of the main challenges in achieving an accurate segmentation result using machine learning approaches. This issue could be addressed using a large training dataset that represents the population diversity. However, due to the lack of sufficient data, we incorporate observer interaction for initializing the CNN to improve the results and make the algorithm more reliable for clinical assessments. We included the user interaction as the second channel of the CNN input. Table 1 shows that for normal subjects, we could achieve test DSC of 92% and 82% for uterine cavity and placenta, respectively, with the absolute volume difference ($|\Delta V|$) of <10% (ΔV was 2% for uterine cavity and -9% for placenta). It is important to note that for uterine cavity even the signed volume difference is close to zero, but the standard deviation of 14% shows that the volume estimation

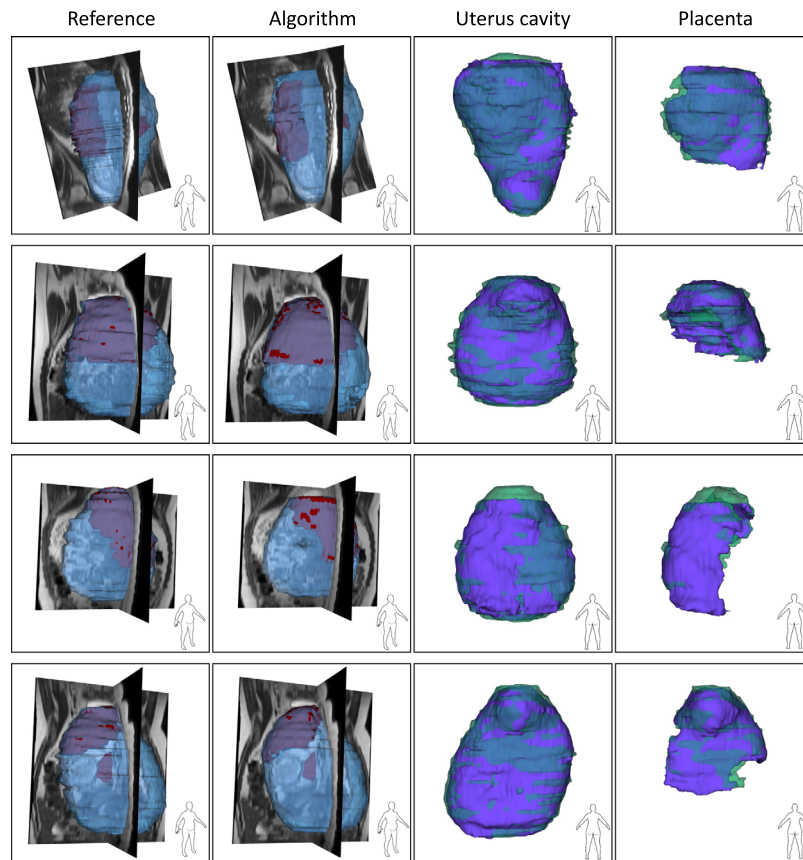


Fig. 6 Qualitative results of the uterus and placenta segmentation in 3D using model II. Each row shows the results for one patient. The two top rows show the results of two subjects from the nonhysterectomy group and the two bottom rows show the results from the hysterectomy group. For the first two columns, the blue, semitransparent shapes show the uterine cavity and the solid, red shapes show the placenta. For the last two columns, the green, semitransparent shapes show the manual segmentations and the solid, purple shapes show the algorithm segmentation results.

accuracy is not consistent across the test data. Therefore, there might be cases with either over- or underestimation of volume. Using other metrics such as DSC and HD compensates the limitation of signed volume difference in evaluating the performance of the algorithm. Using model II trained and tested on both normal and PAS cases did not have a statistically significant impact on the placenta segmentation performance in terms of the DSC metric. However, it significantly reduced the accuracy of the algorithm in uterine cavity segmentation, as determined by the *t*-tests (see Table 2). This could be due to the greater diversity in size and appearance of the uterus after the second dataset with images of abnormal placentas was added. The accuracy of the algorithm in both experiments was substantially higher than the reported results in the literature. The qualitative results in Figs. 4–7 show that the method could segment both the uterine cavity and placenta with acceptable quality. The algorithm detected the shape and position of the uterus and placenta correctly. Therefore, the algorithm could also help detect those abnormalities that are diagnosed based on the relative position of the placenta within the uterine cavity. In some of the results shown in Fig. 4 (e.g., the uterine cavity in the last two rows), there are some small, incorrectly segmented objects seen close to the region of interest that slightly increased the false-positive rate and decreased the measured segmentation accuracy. An extra postprocessing step such as morphological filtering could remove them and slightly increase the accuracy. Comparing the qualitative results to those of previously published work on placenta segmentation^{17,18} shows that our approach outperformed them.

It is important to note that 50% of the images in the second dataset were acquired from women with suspicion of PAS, some of which were severe enough to result in cesarean

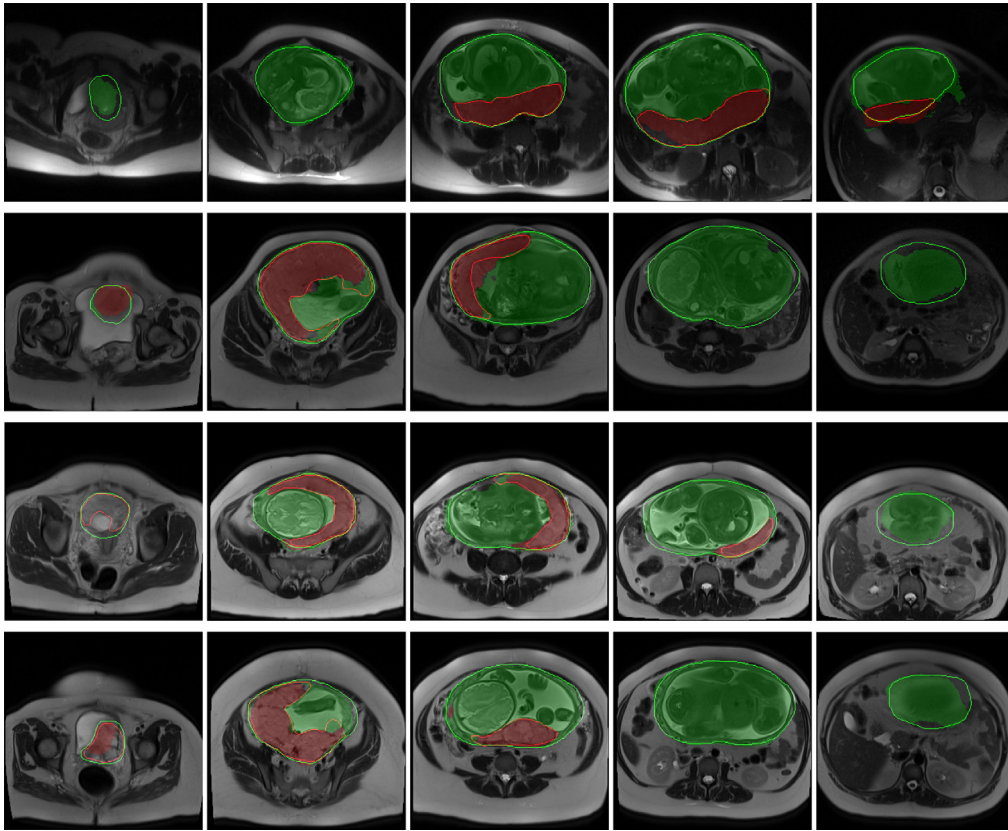


Fig. 7 Qualitative results of the uterus and placenta segmentation in 2D for the four subjects are shown in Fig. 6. Each row shows the results for one patient on five sample axial slices from (left) inferior to (right) superior. The semitransparent green (bright) and red (dark) regions show the algorithm segmentation results for the uterine cavity and placenta, respectively. The solid green (bright) and red (dark) contours show the manual segmentation provided by the expert observer for the uterine cavity and placenta, respectively.

hysterectomy, spanning the spectrum of disease. Because these were performed later in pregnancy, on average, 32 weeks, the average sizes of the uterus and placenta were about 1.5 to 1.8 times larger than those of the first dataset. The images of both datasets were acquired in the same medical center with similar MR scanners and imaging parameters and protocols. Therefore, the observed difference between the results of the first and the second data could be interpreted as the existence of some image-based features that could be considered by a deep learning algorithm for classifying the images to normal and abnormal classes. A comparison between the second and third rows of Table 2 also shows a substantial difference (2% to 4% in terms of DSC and 10% to 14% in terms of ΔV) in placenta and uterine cavity segmentation between hysterectomy and nonhysterectomy groups. In future work, we are going to use a deep learning technique for placenta abnormality prediction and detection.

We used a 3D block-based segmentation approach for better adaptability of the method for image volumes with different numbers of slices. The block-based approach also improved the segmentation accuracy by averaging the output probabilities for the overlapped regions of the blocks.

The average computation time per patient was about 25 s, which is substantially lower than manual segmentation time. However, it is required to measure user interaction time to confirm using this segmentation method could decrease the segmentation time substantially.

4.1 Limitations

The small size of the training set is a limitation. Since there is high intersubject variability in placenta shape and position, the training set must be large enough to represent all the shape and

location variations of the placenta within the uterine cavity. Moreover, the gestational age of the pregnancy affects the fetus's size, appearance, and orientation. This could directly impact the segmentation performance of the algorithm. The large number of training iterations also reflects that issue. Using separate CNN models for each pregnancy trimester could be helpful for more accurate segmentation performance. The long imaging time and the unpredictable and inevitable movements of the baby introduces some imaging artifacts that could have a negative impact on the performance of the segmentation algorithms, especially the 3D segmentation algorithms. An interslice image registration prior to the segmentation could be helpful to mitigate this impact. The other limitation of this work is the block-based segmentation approach. Although it helped with algorithm implementation, it increased the segmentation time and added more computations for integrating the block-level results. The block-based segmentation could improve the segmentation by averaging the probabilities across the blocks overlapped regions. As a future work, we need to employ other CNN architectures that showed better performance in the segmentation of medical images compared to U-Net (e.g., Unet++,²⁶ nnU-Net,²⁷ Unet 3+,²⁸ etc.). We also need to design a user study to test the hypothesis in a more realistic situation, as the current results were achieved based on a simulation of the user interaction.

5 Conclusions

We developed a deep learning approach for simultaneous 3D segmentation of the uterine cavity and placenta in MRI. We guided the network training and testing by incorporating minimal user input for more robust and reliable segmentation performance. We trained and tested the algorithm using two datasets: one dataset with images from normal subjects and the other from a combination of women suspected of PAS, some of whom eventually underwent hysterectomy. The results showed accurate segmentation performance of the algorithm with average DSC values of 92% and 83% for uterine cavity and placenta, respectively. The segmentation method was able to accurately measure placenta size and assess its location, which is an initial step in the application of textural radiomic analysis of the segmented image to aid in the diagnosis of placenta abnormalities.

Disclosures

There is no financial conflict of interest for all the authors.

Acknowledgments

This research was supported in part by the U.S. National Institutes of Health (NIH) (Grant Nos. R01CA156775, R01CA204254, R01HL140325, and R21CA231911) and by the Cancer Prevention and Research Institute of Texas (CPRIT) (Grant No. RP190588). The authors also would like thank Anusha Devi T. T. for assisting with visualization of the results.

References

1. C. M. Zelop et al., "Emergency peripartum hysterectomy," *Am. J. Obstetr. Gynecol.* **168**, 1443–1448 (1993).
2. J. Balayla and H. D. Bondarenko, "Placenta accreta and the risk of adverse maternal and neonatal outcomes," *J. Perinatal Med.* **41**, 141–149 (2013).
3. R. L. Naeye, "Abruptio placentae and placenta previa: frequency, perinatal mortality, and cigarette smoking," *Obstetr. Gynecol.* **55**, 701–704 (1980).
4. W. Sepulveda et al., "Perinatal outcome after prenatal diagnosis of placental chorioangioma," *Obstetr. Gynecol.* **102**, 1028–1033 (2003).
5. C. H. Jansen et al., "Development of placental abnormalities in location and anatomy," *Acta obstetr. Gynecol. Scand.* **99**, 983–993 (2020).
6. R. Gagnon, "Placental insufficiency and its consequences," *Eur. J. Obstetr. Gynecol. Reprod. Biol.* **110**, S99–S107 (2003).

7. A. E. Guttmacher and C. Y. Spong, "The human placenta project: it's time for real time," *Am. J. Obstetr. Gynecol.* **213**, S3–S5 (2015).
8. E. Hafner et al., "Second-trimester measurements of placental volume by three-dimensional ultrasound to predict small-for-gestational-age infants," *Ultrasound Obstetr. Gynecol.* **12**, 97–102 (1998).
9. R. Gassner et al., "Triploidy in a twin pregnancy: small placenta volume as an early sonographical marker," *Prenatal Diagn.* **23**, 16–20 (2003).
10. J. Hasegawa et al., "Placenta volume measurement in clinical practice," *J. Ultrasound Obstetr. Gynecol.* **9**, 408–412 (2015).
11. P. McGinty et al., "Ultrasound assessment of placental function: the effectiveness of placental biometry in a low-risk population as a predictor of a small for gestational age neonate," *Prenatal Diagn.* **32**, 620–626 (2012).
12. J. R. Leyendecker et al., "MRI of pregnancy-related issues: abnormal placentation," *Am. J. Roentgenol.* **198**, 311–320 (2012).
13. C. Maldjian et al., "MRI appearance of placenta percreta and placenta accreta," *Magn. Reson. Imaging* **17**, 965–971 (1999).
14. Q. N. Do et al., "Texture analysis of magnetic resonance images of the human placenta throughout gestation: a feasibility study," *PloS One* **14**, e0211060 (2019).
15. Q. N. Do et al., "MRI of the placenta accreta spectrum (PAS) disorder: radiomics analysis correlates with surgical and pathological outcome," *J. Magn. Reson. Imaging* **51**, 936–946 (2020).
16. S. Dahdouh et al., "In vivo placental MRI shape and textural features predict fetal growth restriction and postnatal outcome," *J. Magn. Reson. Imaging* **47**, 449–458 (2018).
17. G. Wang et al., "Slic-Seg: a minimally interactive segmentation of the placenta from sparse and motion-corrupted fetal MRI in multiple views," *Med. Image Anal.* **34**, 137–147 (2016).
18. A. Alansary et al., "Fast fully automatic segmentation of the human placenta from motion corrupted MRI," *Lect. Notes Comput. Sci.* **9901**, 589–597 (2016).
19. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**, 297–302 (1945).
20. R. Namías et al., "Uterus segmentation in dynamic MRI using LBP texture descriptors," *Proc. SPIE* **9034**, 90343W (2014).
21. Y. Kurata et al., "Automatic segmentation of the uterus on MRI using a convolutional neural network," *Comput. Biol. Med.* **114**, 103438 (2019).
22. M. Shahedi et al., "Segmentation of uterus and placenta in MR images using a fully convolutional neural network," *Proc. SPIE* **11314**, 113141R (2020).
23. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
24. M. D. Zeiler, "ADADELTA: an adaptive learning rate method," arXiv:1212.5701 (2012).
25. M. Abadi et al., "Tensorflow: a system for large-scale machine learning," in *OSDI*, pp. 265–283 (2016).
26. Z. Zhou et al., "UNet++: a nested u-net architecture for medical image segmentation," *Lect. Notes Comput. Sci.* **11045**, 3–11 (2018).
27. F. Isensee et al., "nnU-Net: self-adapting framework for U-Net-based medical image segmentation," arXiv:1809.10486 (2018).
28. H. Huang et al., "UNet 3+: a full-scale connected UNet for medical image segmentation," in *ICASSP 2020-2020 IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, pp. 1055–1059 (2020).

Maysam Shahedi is a research scientist in the Quantitative Bioimaging Laboratory at the University of Texas at Dallas. He received his PhD in biomedical engineering from the University of Western Ontario, Canada. He also received his BSc and MSc degrees in electrical engineering from Isfahan University of Technology. His research interests are medical imaging, medical image processing, image-guided intervention, and machine learning.

Catherine Y. Spong is Professor and Chair of the Department of Obstetrics and Gynecology, Chief of the Division of Maternal-Fetal Medicine, and Paul C. MacDonald Distinguished Chair in Obstetrics and Gynecology. An expert in clinical trials and evidence-based care,

she is particularly focused on optimizing pregnancy outcomes, understanding placentation, inclusion of pregnant and lactating women in research, and eliminating health disparities.

James D. Dormer is the lab manager for the Quantitative Bioimaging Lab at the University of Texas at Dallas, under the direction of Dr. Baowei Fei. His research interests include image-guided surgery, medical device development, and deep learning applications in medicine. He received his bachelor's degree in physics from Franklin and Marshall College and his master's degree in medical physics from the University of Pennsylvania.

Quyen N. Do is an instructor of radiology at UT Southwestern Medical Center and a member of the Radiology Research Division. Her research interests include combinations of innovative multiparametric MR imaging techniques and deep learning-based data analysis to understand human placenta development.

Yin Xi is an assistant professor of radiology at UT Southwestern Medical Center. He received his PhD in statistical science from Southern Methodist University. His research interests include Bayesian clustering, mixture models, image segmentation, and statistical learning.

Matthew A. Lewis is an assistant professor of radiology at UT Southwestern Medical Center. His research interest includes acoustic inverse scattering, spectral CT, biomedical compressive sensing and sparsity, microSPECT and bioluminescent tomography, task-based inverse methods, and ultrasonic assessment of the bone elasticity tensor.

Christina Herrera is an assistant professor of the Department of Obstetrics and Gynecology in the Division of Maternal-Fetal Medicine at UT Southwestern Medical Center. She received her medical degree from Columbia University College of Physicians and Surgeons and completed a residency in obstetrics and gynecology and a fellowship in maternal-fetal medicine at UT Southwestern Medical Center. Her research focuses on noninvasive assessment of the placenta, with the goal to understand pathophysiology of placental-mediated disease via function and development in real time.

Ananth J. Madhuranthakam is an associate professor of radiology at the Advanced Imaging Research Center and Biomedical Engineering Department at UT Southwestern Medical Center. He is the director of magnetic resonance research, and his lab is focused on the technical developments and clinical translation of advanced MR imaging techniques in a variety of applications (lab website: <https://www.utsouthwestern.edu/labs/madhuranthakam/>).

Diane M. Twickler is a professor of radiology with a joint appointment in obstetrics and gynecology at UT Southwestern Medical Center. She is the holder of the Fred Bonte, M.D. Professorship in Radiology and Medical Director of Parkland Obstetrics and Gynecology Ultrasound. She is a fellow of the American College of Radiology and specializes in imaging of obstetrics and gynecology with MR and US. Her research interests include imaging interpretation, postprocessing analytics, and clinical outcomes in fetal dysmorphology and placental abnormalities.

Baowei Fei is a professor of bioengineering and the Cecil H. & Ida Green Chair in Systems Biology Science at the University of Texas at Dallas. He is also a professor of radiology at UT Southwestern Medical Center. He is the Director of the Quantitative Bioimaging Laboratory and Director of the Center for Imaging and Surgical Innovation. He is a fellow of the international society for optics and photonics (SPIE) and the American Institute for Medical and Biological Engineering (AIMBE).