PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Automatic segmentation of uterine cavity and placenta on MR images using deep learning

Shahedi, Maysam, Dormer, James, Do, Quyen, Xi, Yin, Lewis, Matthew, et al.

Maysam Shahedi, James D. Dormer, Quyen N. Do, Yin Xi, Matthew A. Lewis, Christina L. Herrera, Catherine Y. Spong, Ananth J. Madhuranthakam, Diane M. Twickler, Baowei Fei, "Automatic segmentation of uterine cavity and placenta on MR images using deep learning," Proc. SPIE 12036, Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging, 1203611 (4 April 2022); doi: 10.1117/12.2613286

SPIE.

Event: SPIE Medical Imaging, 2022, San Diego, California, United States

Automatic Segmentation of Uterine Cavity and Placenta on MR Images Using Deep Learning

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

[a]Department of Bioengineering, The University of Texas at Dallas, TX
[b]Center for Imaging and Surgical Innovation, The University of Texas at Dallas, TX
[c]Department of Radiology, [d]Department of Clinical Science, [e]Department of Obstetrics and Gynecology, The University of Texas Southwestern Medical Center, Dallas, TX

* Email: bfei@utdallas.edu, Website: https://fei-lab.org

ABSTRACT

Magnetic resonance imaging (MRI) is useful for the detection of abnormalities affecting maternal and fetal health. In this study, we used a fully convolutional neural network for simultaneous segmentation of the uterine cavity and placenta on MR images. We trained the network with MR images of 181 patients, with 157 for training and 24 for validation. The segmentation performance of the algorithm was evaluated using MR images of 60 additional patients that were not involved in training. The average Dice similarity coefficients achieved for the uterine cavity and placenta were 92% and 80%, respectively. The algorithm could estimate the volume of the uterine cavity and placenta with average errors of less than 1.1% compared to manual estimations. Automated segmentation, when incorporated into clinical use, has the potential to quantify, standardize, and improve placental assessment, resulting in improved outcomes for mothers and fetuses.

Keywords: Placenta, uterus, uterine cavity, magnetic resonance imaging (MRI), image segmentation, deep learning

INTRODUCTION

The human placenta is a critical and complex organ that plays a key role in a successful pregnancy. It provides oxygen and nutrition to the growing fetus and vital maternal-fetal exchanges, as well as barriers. Three-dimensional (3D) delineation of the placenta and uterus volumes with magnetic resonance imaging (MRI) has the potential to be useful in studying conditions that result in complications during pregnancy and delivery such as placenta accreta spectrum (PAS), fetal growth restriction, and rarely, intrauterine fetal death[1-4]. Manual segmentation of the placenta MRI to obtain volumes is time-consuming and subject to high inter- and intra-observer variability[5].

As an extension of our previous study on semiautomatic segmentation of the placenta and uterine cavity on MR images[6, 7], here we have expanded to automated segmentation of the placenta and uterine cavity. We used a multi-class segmentation for uterine cavity and placenta segmentation in pregnancy. Previous efforts for developing computerized algorithms for placenta segmentation in MRI[8, 9] presented report low segmentation accuracy with a Dice coefficient [10] of about 72%[9]. Other algorithms need intense computation (*e.g.*, up to two minutes per 3D image volume on average) or require multiple magnetic resonance (MR) image volumes (*e.g.*, sagittal and axial acquisitions) to provide accurate segmentation[8]. Another limitation of the previous study was the small data size[8]. Algorithms presented in the literature for uterine MR segmentation of non-pregnant women[11, 12] have been evaluated as well, when the volumes are smaller than in the 3rd trimester of pregnancy.

Deep learning has demonstrated its strength in fast and relatively accurate segmentation of multiple organ systems from medical images[13, 14]. In this study, our objective is to present an automatic, multi-label segmentation algorithm for fast, accurate, and repeatable 3D segmentation of the uterine cavity and placenta in axial T2-weighted MRI. We used a modified

Medical Imaging 2022: Biomedical Applications in Molecular, Structural, and Functional Imaging, edited by Barjor S. Gimi, Andrzej Krol, Proc. of SPIE Vol. 12036, 1203611
 · 1605-7422 · doi: 10.1117/12.2613286

U-Net architecture as a 3D end-to-end fully convolutional neural network to implement our 3D image segmentation algorithm.

METHODS

2.1 Data

Our dataset contained 241 T2-weighted MR image volumes of the uterus from 241 pregnant women. MR imaging datasets were obtained from both normal and pregnancies with PAS in the second and third trimesters of pregnancy. Each image volume had 28 to 62 two-dimensional (2D) transverse slices. Each slice was originally 256 × 256 pixels in size except for three cases. For those three cases, we resized that image to 256 × 256 by zero-padding to make the slice size consistent across the dataset. The physical size of the image voxels ranged from 1.05 × 1.05 × 7.0 mm^3 to 1.95 × 1.95 × 7.0 mm^3. The labeled images of the uterine cavity and placenta were manually segmented by an expert radiologist.

2.2 Data preparation and preprocessing

For all the images in the dataset, we applied a 2D median filter to all the axial image slices under a 3 × 3 window to reduce the image noise while preserving the edges. Then, to reduce the effect of background voxels and makes the intensity distribution more consistent across the image dataset, we used Equation (1) to truncate and normalize the voxel intensity distribution of each image.

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

where $I_i(x,y,z)$ is the i^{th} image volume in the dataset, and $p_5(I_i)$ and $p_{99.9}(I_i)$ are the 5^{th} and 99.9^{th} percentiles of the voxel intensities in the image volume, respectively. $\hat{I}_i(x,y,z)$ is the normalized image volume.

For each image, there is an overlap between the manual segmentation label of the placenta and that of the uterine cavity. We made the segmentation labels independent from each other using one-hot encoding. We used the portion of the uterus that was not covered by the placenta as the uterine cavity channel and made a three-channel label that included background, uterine cavity (with no overlap with placenta), and placenta labels.

We assumed that the inferior and superior slice of the uterine cavity is known. This could be helpful to limit the size of the data used for training and test. In a clinical setup, defining these slices could be a straightforward process that can be done by the operator. The number of axial slices across the datasets was variable. Therefore, to make the input size consistent across the training and test sets and reduce the size of the network input, we extracted a set of 3D image blocks (B_i^k) from each 3D volume ($\hat{I}_i(x,y,z)$) each contains five sequential 2D axial slices as the 3D input patches. Figure 1 illustrates the image block extraction process from a 3D image volume. For image block extraction, only slices with uterine cavity were selected. For each image, every two adjacent blocks were defined with an overlap of four slices (see Figure 1).

We randomly divided the patients into groups of 157 training (65% of the data), 24 validation (10% of the data), and 60 test cases (25% of the data). Exploiting the left-right symmetry of the images, we used the left-to-right reflection of the image volumes to make two versions of each test image; the original image and the flipped one. We use the average of the output probability maps for both versions to have a more accurate segmentation for each test image.

2.3 Fully convolutional neural network architecture

For this study, we adopted a U-Net [15] architecture which is a fully convolutional neural network (FCNN). We modified the architecture to make it 3D and used that for multi-label image segmentation. Figure 2 shows the network architecture used in this study. The network had four resolution levels with 21 layers including 18 convolutional and three max pool layers in total. We zero-padded the input channels of each convolution/deconvolution layer to keep the size of the output channels the same as the input channels. The input was a 3D image block of the MRI and the output had three channels; one for the background, one for the uterine cavity, and one for the placenta.

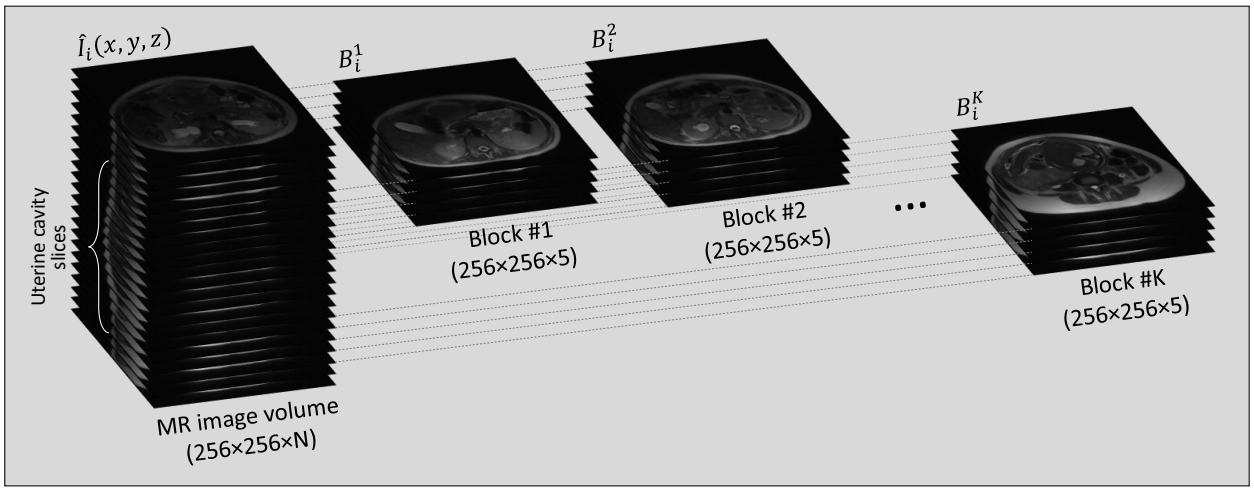


Figure 1. Image blocks extraction from a 3D MR image volume.

We used a Dice similarity coefficient[10] (DSC)-based loss function defined by Equation (2) to address the imbalance between the number of background and foreground (placenta and uterus) voxels:

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

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 mitigate the potential network performance bias toward placenta or uterine cavity segmentation, we calculated two separate loss values, one for the uterine cavity (L_{uterus}) and one for the placenta $(L_{placenta})$ and used the average of the two losses as the total loss. During training, the loss was calculated in block-level. We ran training using the Adadelta[16] gradient-based optimizer for optimization.

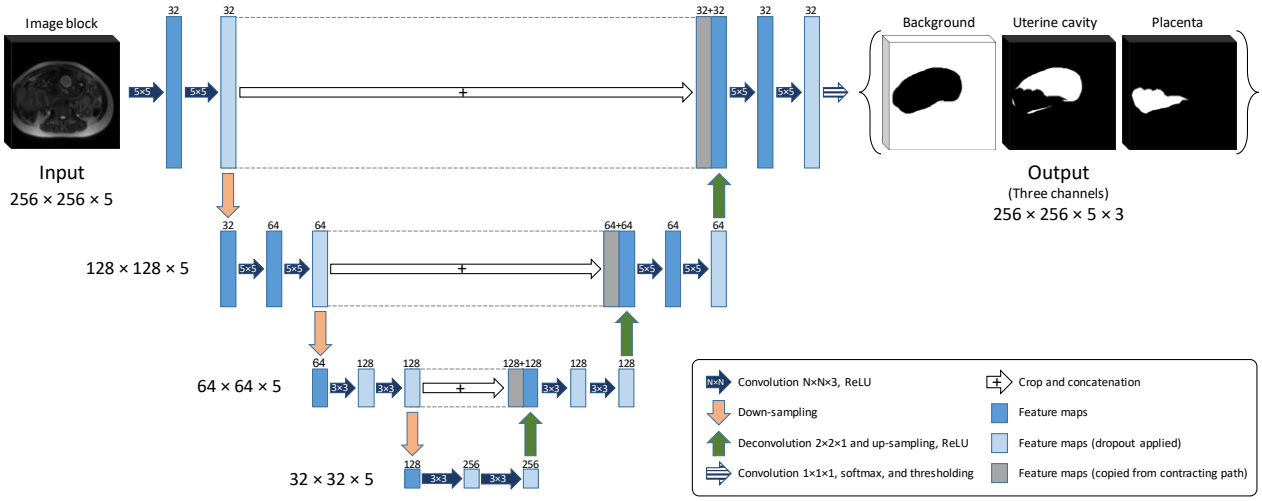


Figure 2. The FCNN architecture (3D U-Net). The number above each layer shows the number of feature maps on that layer. The size of feature maps in each resolution level is mentioned at the lefthand side of the level.

2.4 Implementation details

We implemented the 3D U-Net model in TensorFlow[17] machine learning system using Python platform. We used a high-performance computer with 512 GB of memory and NVIDIA TITAN Xp GPU. We used a batch size of five and an initial learning rate of 1.0, with a dropout rate of 40%, and the decay rate and epsilon conditioning parameter for Adadelta optimizer to 0.9 and 1×10^{-10}.

2.5 Post-processing and evaluation

The network was trained and tested on the 3D image blocks extracted from the MR images. To integrate the block-level results and generate the segmentation result for a 3D test image, we got the voxel-wise average of the block-level probability maps over the overlapped slices and measured a 3D probability map for the whole image volume. We applied the same process to the flipped version of each test image and used the average of the output probabilities corresponding to that image and the flipped image as the final probability map for the test image volume. Thresholding was applied to build output binary masks out of the probability maps. We compared the test results against manual segmentation labels using region-based and volume-based segmentation error metrics including DSC and signed volume difference (ΔV):

$$\Delta V = V_{seg} - V_{ref}, \tag{3}$$

where V_{seg} is the volume of the object of interest (either placenta or uterine cavity) on the algorithm segmentation label and V_{ref} is the volume of the object of interest on the manual segmentation label. We reported DSC in percent and ΔV in cm^3 and percent:

$$\Delta V\ (\%) = \frac{V_{seg}-V_{ref}}{V_{ref}} \times 100. \tag{4}$$

RESULTS

3.1 Training and testing results

We trained the network for up to 500 epochs and chose the model associated with the highest validation accuracy when the validation accuracy trend was plateaued. Figure 3 shows the training and validation accuracy. We reached the highest validation accuracy at epoch number 420 (shown with a dashed line in Figure 3). For the chosen model, the block-level training and validation DSCs were 93.9% and 84.2%, respectively.

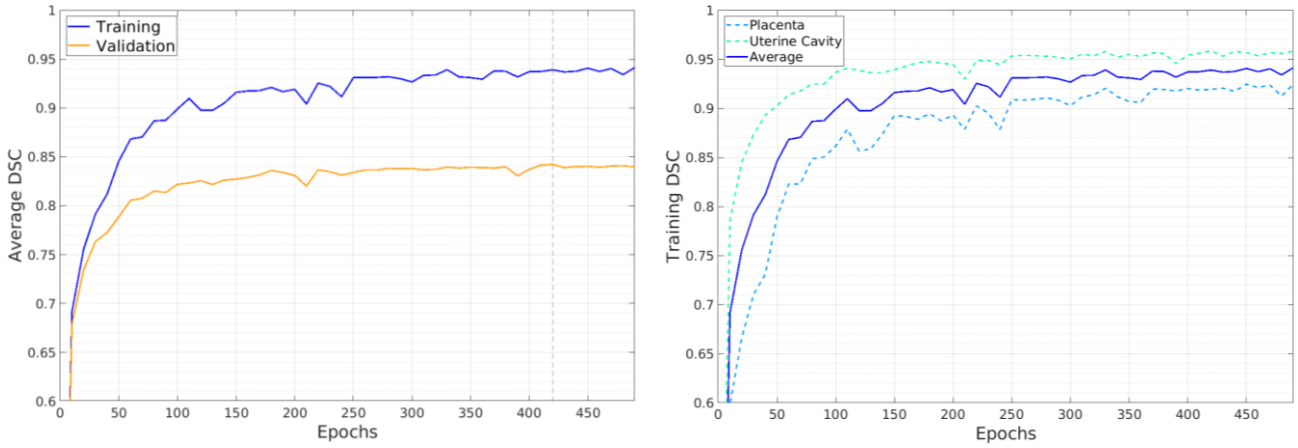
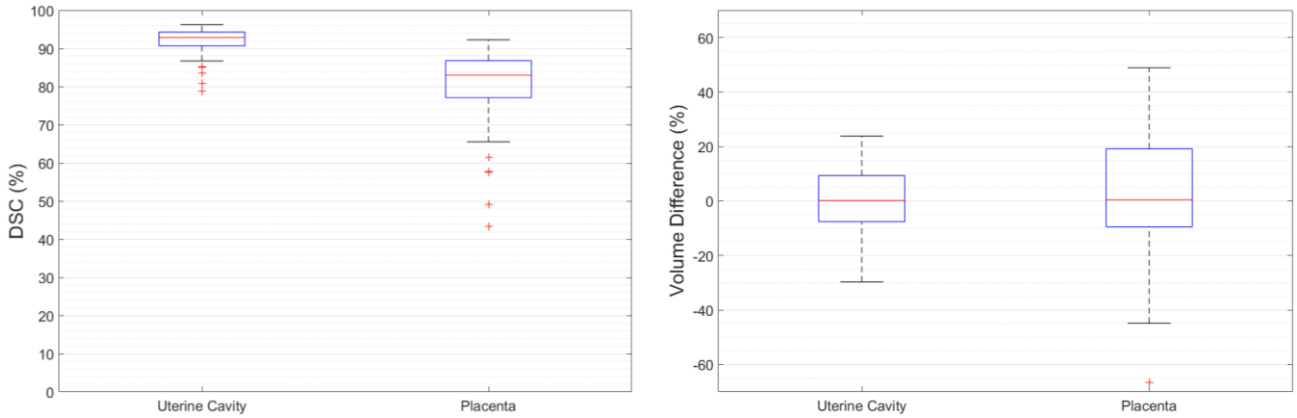


Figure 3. Training results. (a) Block-level training and validation accuracy trends during training based on the average DSCs of the placenta and uterus. The dashed line shows the validation accuracy used to choose the best model. (b) Separate training accuracy trends of the placenta and uterine cavity.

We segmented the test image datasets using the best-trained model and we achieved DSC (mean ± standard deviation) of 91.7 ± 3.7% and 80.2 ± 10.2% for the uterine cavity and placenta, respectively. The boxplot in Figure 4 shows the distribution of the achieved DSC values on the test data for the uterine cavity and placenta. Table I shows the average image-level segmentation error across the test set. The table also compares the results to our previously reported results based on the semi-automatic segmentation algorithm[6]. Using *t*-test with α =0.05, we did not detect any significant difference between automatic and semiautomatic algorithms in terms of DSC. We measured the volume of the uterine cavity and placenta in the test set using the manual segmentation labels provided by the radiologist as a reference. The uterine cavity volume ranged from 1079 cm^3 to 10096 cm^3 and the volume of the placenta ranged from 161 cm^3 to 2239 cm^3. The total computational time for segmenting a 3D image volume was about 14 seconds on average including the integration of the image and its flipped version.



(a) (b)

Figure 4. 3D image-level accuracy of the uterine cavity and placenta segmentation in terms of (a) DSC and (b) ΔV.

Table I. Test results for the uterine cavity and placenta (mean ± standard deviation).

Method	N	Uterine cavity			Placenta		
		DSC (%)	ΔV (cm^3)	ΔV (%)	DSC (%)	ΔV (cm^3)	ΔV (%)
Automatic	60	91.7 ± 3.7	-77 ± 583	0 ± 12	80.2 ± 10.2	20 ± 198	1.1 ± 23
Semiautomatic[7]	50	87.5 ± 5.7	-247 ± 589	-4 ± 14	82.5 ± 5.8	-12 ± 151	0 ± 15

DISCUSSION AND CONCLUSIONS

The proposed multi-class, deep learning segmentation technique could automatically segment the uterine cavity and placenta in 3D MRI volumes with relatively high segmentation accuracy using a single fully convolutional network model. We incorporated minimal observer input (two clicks) for initializing the segmentation with uterine cavity bounding slices. It minimized the search space and made the algorithm more reliable in potential adaptation for clinical use. We also used a 3D block-based segmentation approach for improving the performance (through averaging over the overlapped regions) with a model flexible, employed when the number of axial slices in the input image volumes is variable.

The DSC value for placenta segmentation is lower than that of the uterine cavity because of the higher inter-subject variability of size, shape, location, and MRI appearance of the placenta in comparison with the uterine cavity. However, Figure 4 shows that the average accuracy is affected by a few outliers. For example, for one patient the size of the uterine cavity and placenta (10.1 Liter and 2.2 Liter, respectively) are about two times larger than the largest ones observed in the training set. Figure 4b shows that on average the algorithm could estimate the placenta and uterine cavity accurately, excluding the outliers (median ΔV of about zero for both regions of interest).

In summary, we proposed a fully convolutional deep learning model for automatic segmentation of uterine cavity and placenta in MRI, simultaneously. The method is an expansion of our semiautomatic segmentation algorithm to an automatic deep learning algorithm for simultaneous 3D segmentation of the placenta and uterine cavity in T2-weighted MRI. To the best of our knowledge, our multi-class segmentation model is state-of-the-art in terms of accuracy and speed trade-off as well as clinical adaptability for placenta MR imaging. Comparing the results of the proposed automatic segmentation to our previous study on a semi-automatic approach showed that the algorithm could achieve a similar accuracy level in 3D segmentation of the placenta and uterine cavity with no statistically significant difference. The segmentation algorithm was able to measure placenta size and estimate its location. This first step has important clinical applications which, combined with textural and functional radiomics within the segmentation, will result more quantitiative and reliable evaluation of placenta abnormalities. Future work will focus on automatic initialization and testing the segmentation results in the settings of placental insufficiency and abnormal implantation.

ACKNOWLEDGMENTS

This research was supported in part by the U.S. National Institutes of Health (NIH) grants (R01CA156775, R01CA204254, R01HL140325, and R21CA231911), by the Cancer Prevention and Research Institute of Texas (CPRIT) grant RP190588.

REFERENCES

[1] Leyendecker, J. R., DuBose, M., Hosseinzadeh, K., Stone, R., Gianini, J., Childs, D. D., Snow, A. N., and Mertz, H., “MRI of pregnancy-related issues: abnormal placentation,” American Journal of Roentgenology, 198(2), 311-320 (2012).

[2] Maldjian, C., Adam, R., Pelosi, M., Pelosi III, M., Rudelli, R. D., and Maldjian, J., “MRI appearance of placenta percreta and placenta accreta,” Magnetic resonance imaging, 17(7), 965-971 (1999).

[3] Do, Q. N., Lewis, M. A., Xi, Y., Madhuranthakam, A. J., Happe, S. K., Dashe, J. S., Lenkinski, R. E., Khan, A., and Twickler, D. M., “MRI of the placenta accreta spectrum (PAS) disorder: radiomics analysis correlates with surgical and pathological outcome,” Journal of Magnetic Resonance Imaging, 51(3), 936-946 (2020).

[4] Clark, H. R., Ng, T. W., Khan, A., Happe, S., Dashe, J., Xi, Y., and Twickler, D. M., “Placenta accreta spectrum: correlation of MRI parameters with pathologic and surgical outcomes of high-risk pregnancies,” American Journal of Roentgenology, 214(6), 1417-1423 (2020).

[5] Dahdouh, S., Andescavage, N., Yewale, S., Yarish, A., Lanham, D., Bulas, D., du Plessis, A. J., and Limperopoulos, C., “In vivo placental MRI shape and textural features predict fetal growth restriction and postnatal outcome,” Journal of Magnetic Resonance Imaging, 47(2), 449-458 (2018).

[6] Shahedi, M., Dormer, J. D., TT, A. D., Do, Q. N., Xi, Y., Lewis, M. A., Madhuranthakam, A. J., Twickler, D. M., and Fei, B., "Segmentation of uterus and placenta in MR images using a fully convolutional neural network," SPIE Medical Imaging 2020: Computer-Aided Diagnosis. 11314, 113141R (2020).

[7] Shahedi, M., Spong, C. Y., Dormer, J. D., Do, Q. N., Xi, Y., Lewis, M. A., Herrera, C., Madhuranthakam, A. J., Twickler, D. M., and Fei, B., “Deep learning-based segmentation of the placenta and uterus on MR images,” Journal of Medical Imaging, 8(5), 054001 (2021).

[8] Wang, G., Zuluaga, M. A., Pratt, R., Aertsen, M., Doel, T., Klusmann, M., David, A. L., Deprest, J., Vercauteren, T., and Ourselin, S., “Slic-Seg: A minimally interactive segmentation of the placenta from sparse and motion-corrupted fetal MRI in multiple views,” Medical image analysis, 34, 137-147 (2016).

[9] Alansary, A., Kamnitsas, K., Davidson, A., Khlebnikov, R., Rajchl, M., Malamateniou, C., Rutherford, M., Hajnal, J. V., Glocker, B., and Rueckert, D., "Fast fully automatic segmentation of the human placenta from motion corrupted MRI," International conference on medical image computing and computer-assisted intervention. 589-597 (2016).

[10] Dice, L. R., “Measures of the amount of ecologic association between species,” Ecology, 26(3), 297-302 (1945).

[11] Namías, R., Bellemare, M.-E., Rahim, M., and Pirró, N., "Uterus segmentation in dynamic MRI using lbp texture descriptors," Medical Imaging 2014: Image Processing. 9034, 90343W (2014).

[12] Kurata, Y., Nishio, M., Kido, A., Fujimoto, K., Yakami, M., Isoda, H., and Togashi, K., “Automatic segmentation of the uterus on MRI using a convolutional neural network,” Computers in biology and medicine, 114, 103438 (2019).

[13] Hesamian, M. H., Jia, W., He, X., and Kennedy, P., “Deep learning techniques for medical image segmentation: achievements and challenges,” Journal of digital imaging, 32(4), 582-596 (2019).
[14] Tajbakhsh, N., Jeyaseelan, L., Li, Q., Chiang, J. N., Wu, Z., and Ding, X., “Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation,” Medical Image Analysis, 63, 101693 (2020).
[15] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," International Conference on Medical image computing and computer-assisted intervention. 234-241 (2015).
[16] Zeiler, M. D., “Adadelta: an adaptive learning rate method,” arXiv preprint arXiv:1212.5701, (2012).
[17] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., and Isard, M., "Tensorflow: A system for large-scale machine learning," 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16). 265-283 (2016).