

Received November 13, 2019, accepted November 28, 2019, date of publication December 6, 2019, date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958133

# Automatic Segmentation of Human Placenta Images With U-Net

MO HAN<sup>1</sup>, YUWEI BAO<sup>2</sup>, ZIYAN SUN<sup>2</sup>, SHIPING WEN<sup>1</sup>, LIMING XIA<sup>2</sup>, JINGYANG ZHAO<sup>3</sup>, JUNFENG DU<sup>4</sup>, AND ZHENG YAN<sup>5</sup>

<sup>1</sup>School of Computer Science and Engineering, University of Electronics Science and Technology of China, Chengdu 611731, China

<sup>2</sup>Department of Radiology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China

<sup>3</sup>School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>4</sup>Department of Plastic Surgery, Liyuan Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430077, China

<sup>5</sup>University of Technology Sydney, Ultimo, NSW 2007, Australia

Corresponding authors: Ziyang Sun (ziyansun@hust.edu.cn) and Shiping Wen (wenshiping@uestc.edu.cn)

This work was supported by the Natural Science Foundation of China under Grant 61673187.

**ABSTRACT** Placenta is closely related to the health of the fetus. Abnormal placental function will affect the normal development of the fetus, and in severe cases, even endanger the life of the fetus. Therefore, accurate and quantitative evaluation of placenta has important clinical significance. It is a common method to segment human placenta with semantic segmentation. However, manual segmentation relies too much on the professional knowledge and clinical experience of the staff, and it will also consume a lot of time. Therefore, based on u-net, we propose an automatic segmentation method of human placenta, which reduces manual intervention and greatly speeds up the segmentation, making large-scale segmentation possible. The human placenta data set we used was labeled by experts, which was obtained from prenatal examinations of 11 pregnant women, about 1,110 images. It was a comprehensive and clinically significant data set. By training the network with such data set, the robustness of the model will be better. After testing on the data set, the segmentation effect is basically consistent with the manual segmentation effect.

**INDEX TERMS** Automatic segmentation, human placenta image, semantic segmentation, U-net.

## I. INTRODUCTION

Placenta is an important organ for the exchange of substances between fetus and mother, which plays a vital role in the healthy growth of fetus. The fetus develops in the womb and depends on the placenta for its nourishment. What's more, the placenta also synthesizes many substances that maintain pregnancy, including estrogen, progesterone, cytokines, growth factors and so on. In addition, the placenta also has the function of protecting the fetus, and it has a certain barrier function against some bacteria, pathogens and drugs, though its function is very limited. Placental lesions can have serious consequences, such as intrauterine growth restriction (IUGR) due to placental insufficiency, which reduces the growth potential of the fetus and increases the risk of uncertainty in subsequent growth [1]. In the case of twins, abnormal blood vessels in the placenta can cause twin-to-twin transfusion syndrome (TTTS), resulting in unbalanced blood distribution between twins, which can lead fetal to death in

severe cases [2]. Placental hyperplasia will increase the risk of heavy bleeding during delivery [3]. Therefore, the placenta is closely related to the development of the fetus. Through the examination of the placenta, the health information of the fetus and the mother can be obtained. This may assist the physician in making a diagnosis. It is a common method to detect the placenta to obtain images of the abdomen by Nuclear Magnetic Resonance Imaging (MRI) and then segment the placenta in the abdomen by relevant algorithms. The position of the fetus in the mother's body is not fixed, but changes with time. The breathing movement of the mother causes the position of the fetus to change, and the placenta to move accordingly. Inevitably, this also leads to changes in the shape of the placenta. On the other hand, the size and shape of the placenta change as the fetus develops. This series of position and shape changes will cause difficulties in segmentation.

As one of the important topics in computer vision, Semantic segmentation is widely used in the fields of human-computer interaction [4], autonomous driving [5] and medical image processing [6]. Deep learning algorithms, especially

The associate editor coordinating the review of this manuscript and approving it for publication was Shirui Pan.

convolutional neural networks [7]–[23], have achieved the best results in the field of computer vision. Convolutional network is well known for its excellent performance in image classification, however, the concept has been proposed as early as 1989 by LeCun *et al.* [24]. Due to the limited computing power and the lack of data at that time, convolutional network failed to achieve the desired effect. Krizhevsky *et al.* [25] improved on predecessors and trained a network with a large number of parameters on ImageNet. With the continuous optimization of the network by researchers, the network has more and more layers, and the classification effect is getting better and better [26]–[55]. The traditional convolutional neural network is mainly used for classification. The output is the probability of each object, and decides which category the input object belongs to according to the probability. However, in many cases, we need not only to categorize the goals, but also to understand the deeper information. For example, if we successfully classify a picture of a cat, but we also need to know a certain part of the cat, such as the eyes, then the output of the network is not the probability of the category, but the whole image. Therefore, Long *et al.* [56] made improvements on the basis of convolutional network and replaced the fully connected network with convolutional layer in the network, achieving good results in image segmentation. Since the fully convolutional network was proposed, it has occupied an important position in the field of semantic segmentation. Many researchers have made improvements on the basis of the fully convolutional network in an attempt to achieve better results. Among them, U-net is an excellent one, and it shines in the field of medical image segmentation. Compared with the segmentation in other fields, the semantic segmentation precision of medical image is higher, and the data sets in the medical field are relatively scarce, as well as the capacity of data sets is relatively small. The appearance of U-net solves this problem better. In this paper, U-net is used to segment the MRI placenta image, and good results are obtained.

## II. RELATED WORK

The placenta plays an important role in the healthy development of the fetus. Therefore, many researchers have proposed methods to segment the placenta. The following is a brief introduction to the past work.

Wang *et al.* [57] proposed an interactive learning based semi-automatic segmentation method, called Slic-Seg, which firstly used random forest to train labeled data and formed a slice by slice framework to realize the segmentation of placenta in MRI images. Stevenson *et al.* [58] combined the original random walk method with virtual organ computer-aided analysis (VOCAL) to segment the placenta in 3D ultrasound images. Wang *et al.* [59] improved the Slic-Seg method in 2016, and proposed a probability-based 4D Graph Cuts method, which used the consistency between inter-slice and inter-image to refine the segmentation, achieving better segmentation effect. Alansary *et al.* [60] proposes a fully automatic segmentation framework used in MRI image

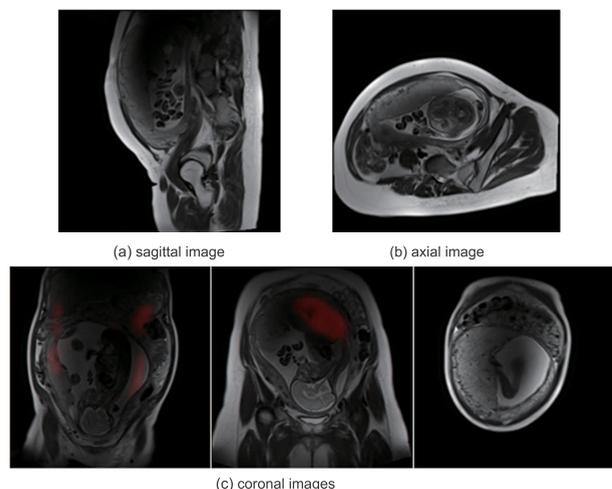
segmentation of human placenta, they firstly use 3D multi-scale convolution neural network to identify the placenta candidate area, and then use 3D dense conditional random field to refine the results. They test this approach in 66 pregnant women, relatively good results have been achieved. Looney *et al.* [61] combined deep convolutional neural network with random walk algorithm, and used 300 ultrasonic images for training, verification and testing, showing that the convolutional neural network can automatically segment the placenta in 3D ultrasound images.

The visualization of the placenta is important for the assessment of placental health. Miao *et al.* [62] used advanced motion compensation and automatic segmentation to extract the shape of placenta, and extracted a visualized technique to display the fetal side and maternal side of placenta, which provided help for doctors to analyze pathological information. Yang *et al.* [63] proposed a framework for simultaneously segmenting fetus, gestational sac and placenta in 3D ultrasound images. They added recursive neural network on the basis of 3D fully convolutional network to refine local segmentation results from the perspective of semantic information. In addition, the authors also introduced the hierarchical deep supervision mechanism which promoted the communication of information within the network and further improved the result of semantic segmentation. They tested this method on their dataset and achieved good segmentation results. Gibson *et al.* proposed an algorithm for abdominal multi-organ segmentation based on deep learning, the organs included pancreas, gastrointestinal tract (esophagus, stomach, duodenum), liver, spleen, left kidney and gallbladder. The algorithm is verified in a dataset composed of more than 90 subjects, and the segmentation effect is excellent.

In 2018, Wang *et al.* [64] proposed a deep learning-based interactive segmentation method to improve the segmentation results of convolutional neural network. Firstly, the author used a convolutional neural network for initial segmentation, and added user interactions to mark the error segmentation. Then, the results of the first neural network were processed with the second convolutional neural network to further improve the precision and effect of segmentation. The effectiveness of the proposed algorithm was verified by 2D MRI image segmentation of the placenta. Looney *et al.* [65] proposed a network to segment placenta from 3D ultrasound images, called OxNNNet (a fully convolutional network). The obtained results are basically consistent with the clinical results, however, the network has requirements for datasets. The larger the datasets, the better the performance.

## III. METHODS

The semantic segmentation network is used for image segmentation of the placenta obtained by MRI, and the accurate segmentation of the placenta is achieved by this method. We also validate the impact of different hyperparameters on network accuracy. In this way, accurate segmentation of the placenta is achieved.



**FIGURE 1.** Three types of placental images. (a) The sagittal image of placentas; (b) The cross-sectional image of placentas; (c) The coronal images of placentas.

### A. DATA

All MRI images of 11 normal placentas were labeled with placental contour, including 425 axial images, 271 coronal images and 414 sagittal images. Images were collected using a 3.0T magnetic resonance scanner. Pregnant women were in supine position or left lateral position, and the placenta was used as the center to scan the three planes of placenta (transverse plane, sagittal plane and coronal plane). Labeling was performed by a physician with three years of experience in placental imaging diagnosis using software to perform multipoint tracking labeling (the distance between points was less than 5mm). The placenta's overall attachment position, size and shape were observed on the transverse axis, coronal and sagittal planes of the placenta before labeling, and then manually labeled layer by layer. Each placental marker includes two parts:

1. Placental attachment surface: the mark points strictly follow the deformation of basal decidua between placenta and uterus, a small part of the local basal decidua shows unclear, so it can be reasonably delineated according to the upper and lower layers and the multidirectional contour of placenta.

2. Uterine surface of placenta: the marked points are strictly on the boundary between placenta and amniotic fluid. When the contour of placenta shows more tortuosity, the number of marked points will be increased appropriately.

Axial, coronal and sagittal images of the embryo were taken from different directions and angles. The analysis of the images from different directions is conducive to the fine segmentation of the placenta. As we all know, medical data sets are usually difficult to obtain. Our placental data set comes from the prenatal examination of 11 pregnant women, each of which took about 100 images and was labeled by professional doctors. Our data set is comprehensive and consistent with the reality. Using such a data set to train the

network, the obtained model also has strong robustness and good clinical significance, which is also a contribution of this paper.

### B. DATA FORMATTING

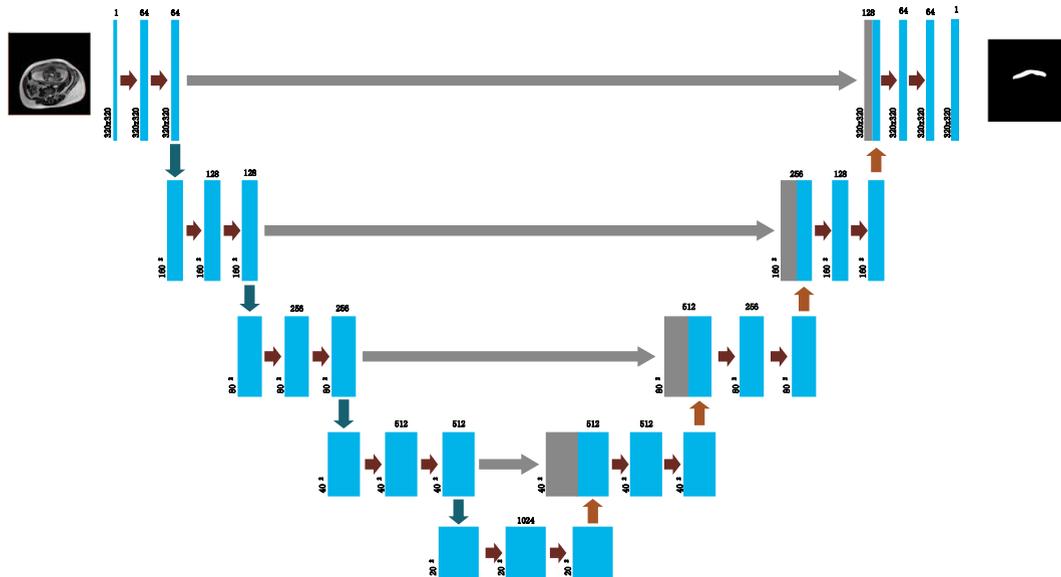
In semantic segmentation and other tasks, preprocessing data is an important part. Because the original data distribution may vary greatly or not belong to the same distribution. In order to fully utilize the performance of the network and improve the accuracy of the network, different preprocessing methods are adopted for different tasks. In object detection, feature normalization and random flipping are effective. It is useful to normalize data in semantic segmentation. However, for data sets with less data, using random flipping can effectively increase the diversity of data sets.

We made the image adapt well to this network through several methods of data preprocessing. Since the picture obtained by MRI is a cross section of a part of the abdomen and is intercepted at different parts, the difference in the picture is relatively large. Obtained from the MRI device is a black and white picture, so for the convenience of observation, we turn black and white into a color picture. Since we get a large picture, in order to reduce the need for computation and memory, we have all the pictures into  $[320,320]$  size. In order for the network to process this data accurately, we need to normalize the image to  $[0,1]$ .

### C. SEMANTIC SEGMENTATION

When semantic segmentation is proposed, the main problem is to split the position of the object in the image. However, with the development of other fields, semantic segmentation also plays an important role, such as automatic driving, medical image analysis, etc., which depend on the development of semantic segmentation techniques. Although object detection can also determine the position of an object, semantic segmentation is pixel-level and the results are more accurate. This paper applies semantic segmentation techniques to medical image analysis.

Placental segmentation is to segment the position of the placenta in the MRI image, which is essentially semantic segmentation. Semantic segmentation is the classification of each pixel in an image to segment objects with specific features. We enter the original MRI image into the network and train with the already-made tags. The label indicates the position of the placenta in the original image. The output of the network is also a black and white picture, and the white color indicates the position of the placenta. We use the structure of encoding-decoding, which is typically U-net. We extract the features of the picture step by step through the convolutional layer, and the features are from low-dimensional to high-dimensional, and through the combination of decoding and low-dimensional features. There are also long-distance connections in the network. In this way, the low-dimensional features can be merged in the upsampling process, thus the network can capture the global and detailed information well, which is beneficial for each pixel classification.

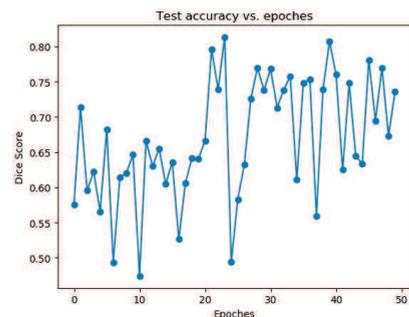


**FIGURE 2.** U-net is used for placental segmentation. A captured picture is entered into the network and processed to obtain an image that separates the placenta. It can be seen that the entire network is a process of encoding to decoding. Through such a method, we can effectively combine low-dimensional and high-dimensional features to make the segmentation result more accurate.

#### D. U-NET

The network architecture we use is U-net. The effectiveness of this network structure has achieved good results in the field of semantic segmentation. Such a network structure can simultaneously utilize details and global information to increase the accuracy of segmentation. Although the network has downsampling and upsampling operations, it is through such operations that the network obtains global information. For the details, we combine the details and the global information through a cross-layer connection. In this way, the network can accurately find the position of the placenta in images. We preserved the four upsampling and downsampling layers in the original structure because the validity of this structure has been verified in semantic segmentation. Since we input a black and white image, the number of layers of the first convolutional layer for input becomes 1. The final result output is also one layer, because we only need to segment the position of the placenta. So the two parts of the network have changed. At the same time, the network requires less memory capacity, which can fully utilize the performance of the machine and shorten the training time. A number of papers have already demonstrated the effectiveness of this structure. Although U-net is an improvement based on FCN, it is not a simple way of encoding to decoding. The resulting high-dimensional features are merged with the features of the lower layers during the upsampling process to preserve the details of the image to the greatest extent possible.

U-net is used for placental segmentation. A captured picture is entered into the network and processed to obtain an image that separates the placenta. It can be seen that the entire network is a process of encoding to decoding. Through such

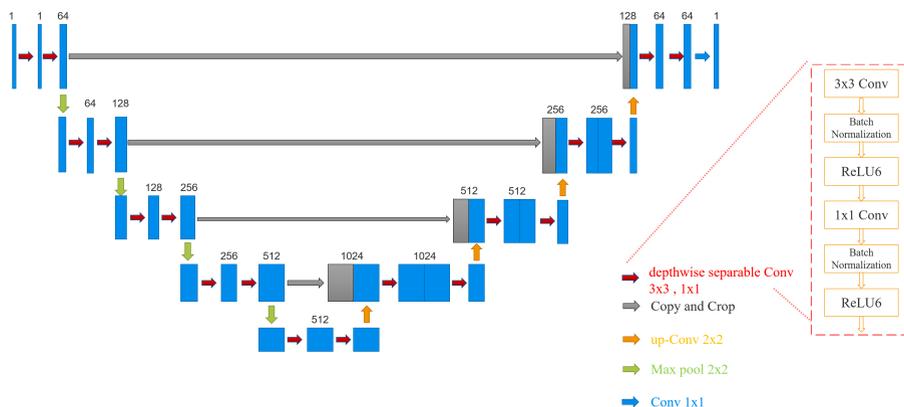


**FIGURE 3.** The relationship between accuracy and training times.

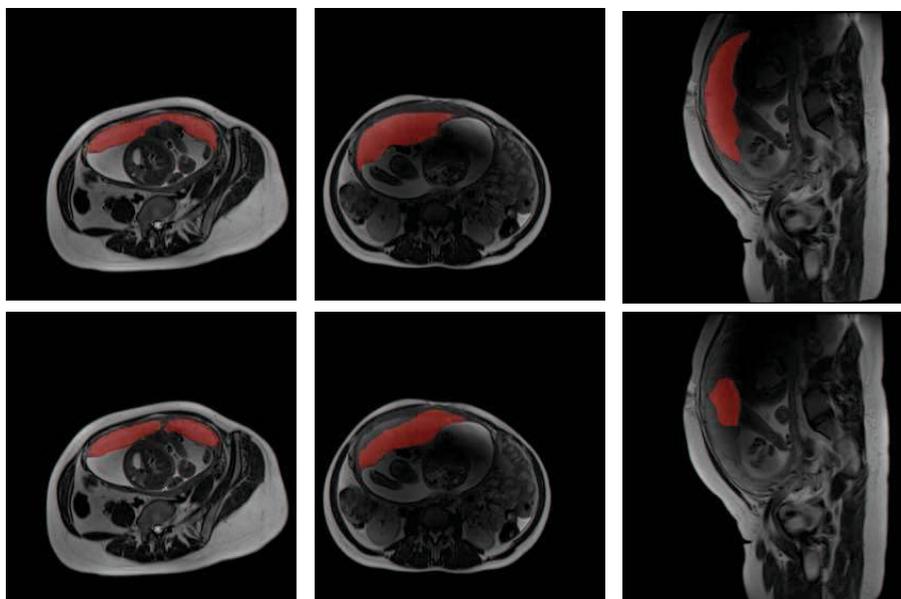
a method, we can effectively combine low-dimensional and high-dimensional features to make the segmentation result more accurate. And in order to reduce the amount of calculation, there is no fully connected layers in the network, so that better segmentation results can be achieved with a minimum of parameter quantities. The downsampling process is followed by every  $2 \times 3$  convolutional layer followed by a  $2 \times 2$  maximum pooling layer. The ReLU activation function is used after each convolutional layer, and the number of channels is gradually increased during the downsampling process. In the upsampling process, each step consists of a  $2 \times 2$  convolutional layer and two  $3 \times 3$  convolutional layers, and the activation function is also ReLU. In the last layer, a  $1 \times 1$  convolutional layer is used to turn the final output into an image with channel 1.

#### E. SETTINGS AND EVALUATION

All experiments were performed on Pytorch. The optimizer we use is SGD and the learning rate is set to 0.01. The batch



**FIGURE 4.** Separable U-net architecture with convolutional encoder and decoder using separable convolution based on U-net architecture, It can construct the model thinner and less computationally expensive.



**FIGURE 5.** Forecast performance display. The above is ground truth, the following is the prediction result.

size is 8, which can effectively increase the training efficiency. For the initialization of the ownership heavy parameters, we use random initialization. The cross-entropy loss function is used to calculate the loss, because this method can be used to distinguish categories well.

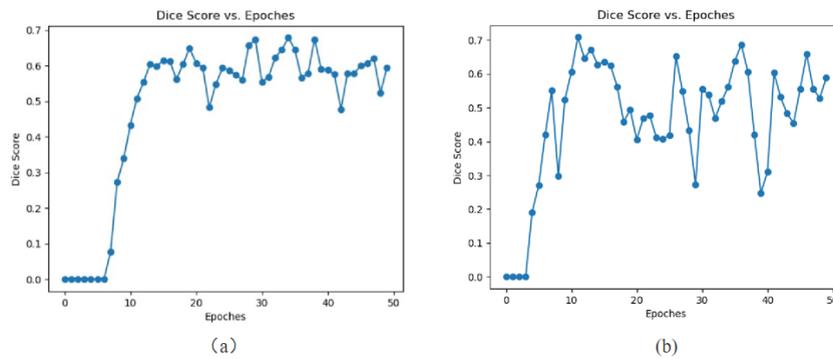
The U-net network trained 50 rounds on Titan X. We verified each training and marked the dice score. As you can see in Fig. 3, for a relatively simple data set such as placental segmentation, U-net is able to converge quickly and achieve good accuracy. There are two reasons for having such an effect. The convolutional neural network can well learn the characteristics of the picture and fuse the local and global information. The second is that for the task of placental segmentation, only the second classification is performed, so the task is relatively simple for the U-net network.

We show a split graph for each of the different directions and positions. Since MRI images are acquired at different locations when acquired. So the images obtained are also

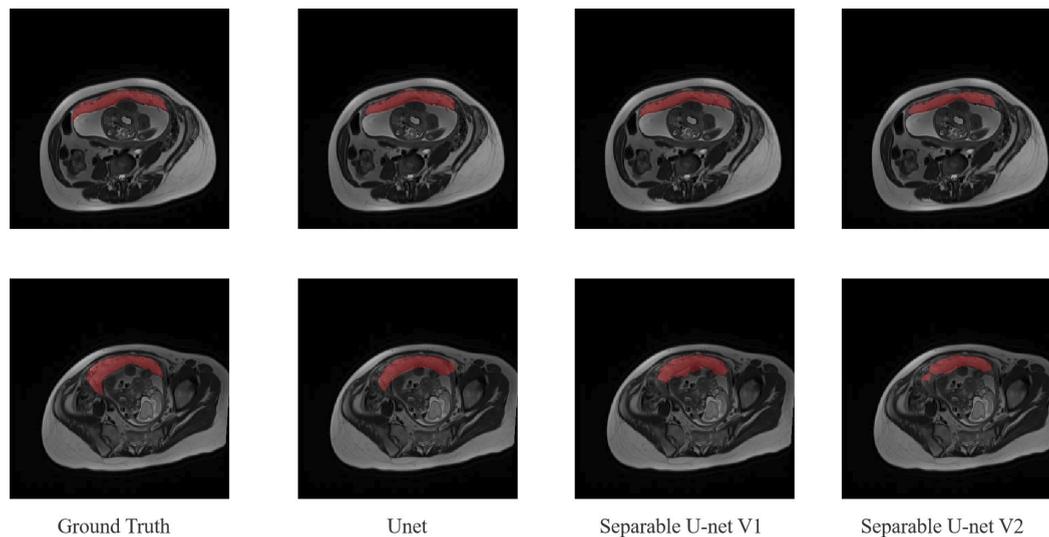
different. To illustrate whether the acquired position and orientation have an effect on the segmentation results, we have shown the results of each segmentation. As can be seen from Fig. 5, the influence of this factor on the segmentation result is relatively small.

#### F. SEPARABLE U-NET

As we all know, deep neural networks have evolved to the state-of-the-art technique for computer vision tasks [27]. Unfortunately, neural networks are not only computationally intensive but also memory intensive, making them hard to deploy on mobile and embedded system. If U-net can be deployed on mobile medical devices, it will make it easier for doctors to diagnose whether the placental is abnormal or not. Inspired by U-net and MobileNets [66], we proposed a network named Separable U-net for placental segmentation task. The network we proposed also has an overall architecture



**FIGURE 6.** (a): Separable U-net V1 relationship between accuracy and training times; (b): Separable U-net V2 relationship between accuracy and training times;



**FIGURE 7.** Forecast performance display. The first column shows the ground truth, the second and third, fourth column show the predicted result obtain with U-net, Separable U-netV1, V2.

similar to the standard U-net, it has encode function on the left side of architecture and right side for decode function which has thinner model size so it can be easily deployed on mobile and embedded system. Fig. 4 illustrates the network architecture. The standard convolution was replaced by the depthwise separable convolution [67], detail design is shown in the dashed window which consisted of a  $3 \times 3$  convolutional layer called depthwise convolution, a batch normalization layer [68] and an activation layer,  $1 \times 1$  convolution for pointwise convolution.

Compared with traditional U-net architecture, which we introduced above, the network we proposed achieve lower computational cost and reduce the number of parameters of U-net model.

#### IV. EXPERIMENTAL RESULTS

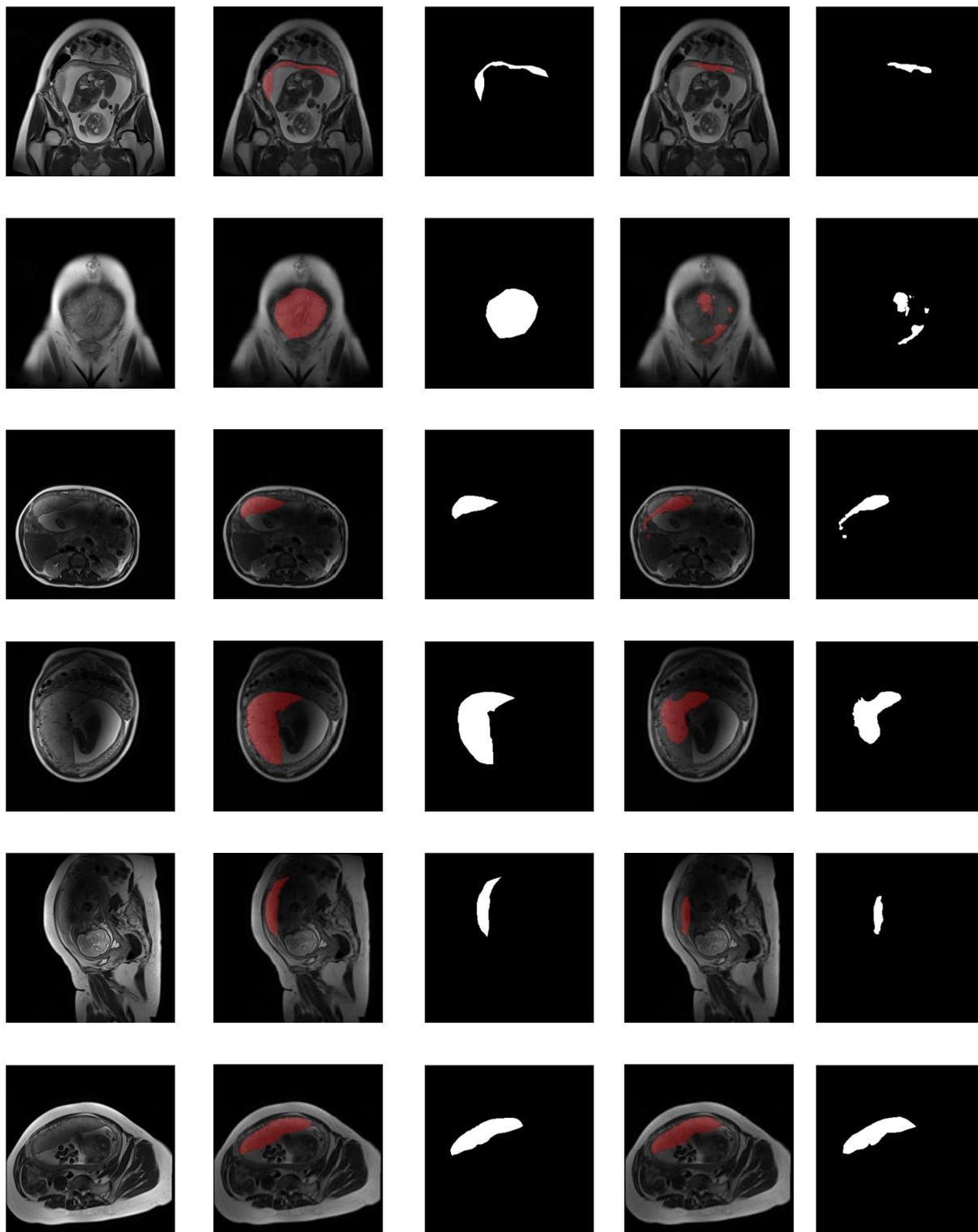
We use several different criteria to analyze the accuracy of the network. We use the overall accuracy, mean accuracy and mean IU indicators to measure the accuracy of the network. Pixel Accuracy is the simplest metric for marking the correct pixel count as a percentage of the total pixels. Mean accuracy

is a simple improvement of PA that calculates the proportion of pixels that are correctly classified within each class, and then averages all classes. Mean IU is the standard measure of semantic segmentation. It calculates the intersection of the two sets and the union of the two sets. In the problem of semantic segmentation, the two sets are real values and predicted values. It can be seen from the comparison of the accuracy that the network can already segment the placenta well.

As can be seen from Fig. 8, our network can effectively segment the placenta. But for some difficult pictures, only the approximate position can be split. Therefore, our network is very effective for placental segmentation.

##### A. COMPARISON WITH SEPARABLE U-NET AND U-NET

In order to reduce the model size and find out the relationship between accuracy and the number of convolutional channels, we design two types of Separable U-net with different channels, both of them and traditional U-net use the same learning rate, batch size and optimizer. Also, we initialized the weights with random values, set the batch size to 10 and trained



**FIGURE 8.** Examples of segmentations obtain with U-net, showing the effect of the proposed network. The first column shows the original images, the second and third columns show the ground truth, the fourth and fifth columns show the predicted results.

50 rounds on NVIDIA GeForce GTX 1080Ti GPU, SGD as optimizer and binary cross-entropy as our loss function.

From Table 1, to distinguish two types of Separable U-net, we named them Separable U-net V1 and V2. It shows that

U-net achieves the highest accuracy and mean IU, but Separable U-net has the lowest memory cost with a little accuracy loss. Compared with traditional U-net, it may more useful in mobile and embedded system. Fig. 6 performs two

**TABLE 1. Performance of the three models on placental data sets.**

method	overall accuracy	mean accuracy	mean IU	model size
U-net	<b>0.9868</b>	<b>0.8671</b>	<b>0.8171</b>	53.6MB
Separable U-net V1	0.9799	0.8445	0.7568	29.7MB
Separable U-net V2	0.9805	0.8549	0.7643	<b>19.3MB</b>

types of Separable U-net's relationship between accuracy and training times. V1's accuracy is more stable than V2 during the training process and V2's Maximum accuracy is larger than V1. Fig. 7 shows the predicted pictures obtain with U-net, Separable U-net and ground truth. From the pictures we can conclude that Separable U-net we proposed can also effectively segment the placenta.

### B. COMPARISON WITH EXISTING METHODS

In order to reflect the advantages of U-net in placental segmentation, We also compared our method with several existing methods, FCN and Deeplab, which are standard segmentation algorithms, Table 2 shows that different type of methods' performance. From the results, it shows that U-net performs the best among those methods. It achieves the highest overall accuracy of 0.9868, the highest mean accuracy of 0.8671 and the highest mean IU of 0.8171 with a small quantity of dataset, which shows that the U-net exhibits state-of-the-art performance comparing both standard segmentation methods. Moreover, U-net has the smallest model size so it can be deployed on mobile system easily. To summarize, U-net performs best in placental segmentation.

**TABLE 2. Placenta segmentation results.**

method	overall accuracy	mean accuracy	mean IU	model size
U-net	<b>0.9868</b>	<b>0.8671</b>	<b>0.8171</b>	<b>53.6MB</b>
FCN	0.9546	0.7711	0.6236	537.1MB
Deeplab	0.9779	0.7024	0.6777	239.7MB

### V. DISCUSSION

According to the results of the final segmentation, it is very effective to use U-net to segment simple or relatively straight placentas. Although the segmentation results for the relatively elongated placenta are not very good, the body portion of the placenta is clearly segmented. As well known, it is very time-consuming to distinguish the placenta by the naked eye to observe the images. The use of convolutional neural networks can process data in a short period of time and produce a more accurate result. Therefore, applying U-net to medical problems can reduce costs and improve medical efficiency. The data we get is not only from different people, but also includes MRI images of the same person in different parts. This makes the network encounter a relatively large problem in the learning process, because the proportion of the entire image occupied by the placenta in some parts is relatively large, it will be easier to segment, and the accuracy will

be better. For some pictures with a small proportion, the placenta is too small to be accurately segmented. Therefore, this problem will make the positive and negative samples unbalanced, which will cause a certain deviation in the training of the network. This is an important factor affecting the accuracy of the network.

### VI. CONCLUSION

In summary, experiments have shown that the use of U-net can be a good way to segment the placenta. Based on the results of our final test, we can see that though some images are not completely segmented, their approximate location is already observable. For simple pictures, the placenta can be easily segmented. For the segmentation results, you can see that the picture with a larger proportion of the placenta is better. For a relatively small picture, the segmentation result will be inconsistent. Although some results are not ideal, verification can prove that the placenta can be well segmented through U-net. In addition, through the quantified results, we can see that U-net can segment the placenta very well. A large amount of medical cost can be saved by such an approach. Moreover, the reduction of human intervention makes the segmentation speed faster and enables large-scale segmentation operations.

### CONFLICT OF INTEREST STATEMENT

There is no conflict of interest in this paper.

### ACKNOWLEDGMENT

(*Mo Han and Yuwei Bao are co-first authors.*)

### REFERENCES

- [1] A. A. Baschat, "Fetal responses to placental insufficiency: An update," *BJOG: Int. J. Obstetrics Gynaecol.*, vol. 111, no. 10, pp. 1031–1041, 2004.
- [2] J. A. Deprest, A. W. Flake, E. Gratacos, Y. Ville, K. Hecher, K. Nicolaidis, M. P. Johnson, F. I. Luks, N. S. Adzick, and M. R. Harrison, "The making of fetal surgery," *Prenatal diagnosis*, vol. 30, no. 7, pp. 653–667, 2010.
- [3] C. Mazouni, G. Gorincour, V. Juhan, and F. Bretelle, "Placenta accreta: A review of current advances in prenatal diagnosis," *Placenta*, vol. 28, no. 7, pp. 599–603, 2007.
- [4] M. Oberweger, P. Wohlhart, and V. Lepetit, "Hands deep in deep learning for hand pose estimation," 2015, *arXiv:1502.06807*. [Online]. Available: <https://arxiv.org/abs/1502.06807>
- [5] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 3213–3223.
- [6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Springer, 2015, pp. 234–241.
- [7] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1717–1724.
- [8] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3320–3328.
- [9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [10] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, and M. Bernstein, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

- [11] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: When to warp?" in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov./Dec. 2016, pp. 1–6.
- [12] Z. Yan, W. Liu, S. Wen, and Y. Yang, "Multi-label image classification by feature attention network," *IEEE Access*, vol. 7, pp. 98005–98013, 2019.
- [13] S. Wen, W. Liu, Y. Yang, T. Huang, and Z. Zeng, "Generating realistic videos from keyframes with concatenated GANs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 8, pp. 2337–2348, Aug. 2019.
- [14] G. Ren, Y. Cao, S. Wen, Z. Zeng, and T. Huang, "A modified Elman neural network with a new learning rate scheme," *Neurocomputing*, vol. 286, pp. 11–18, Apr. 2018.
- [15] M. Dong, S. Wen, Z. Zeng, Z. Yan, and T. Huang, "Sparse fully convolutional network for face labeling," *Neurocomputing*, vol. 331, pp. 465–472, Feb. 2019.
- [16] Z. Li, M. Dong, S. Wen, X. Hu, P. Zhou, and Z. Zeng, "CLU-CNNs: Object detection for medical images," *Neurocomputing*, vol. 350, pp. 53–59, Jul. 2019.
- [17] S. Wen, H. Wei, Z. Yan, Z. Guo, Y. Yang, T. Huang, and Y. Chen, "Memristor-based design of sparse compact convolutional neural network," *IEEE Trans. Netw. Sci. Eng.*, to be published.
- [18] S. Wen, H. Wei, Y. Yang, Z. Guo, Z. Zeng, T. Huang, and Y. Chen, "Memristive LSTM network for sentiment analysis," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published.
- [19] X. Xie, L. Zou, S. Wen, T. Huang, and Z. Zeng, "A flux-controlled logarithmic memristor model and equivalent circuit," *Circuits, Syst., Signal Process.*, vol. 38, no. 5, pp. 1452–1465, 2019.
- [20] S. Wen, S. Xiao, Y. Yang, Z. Yan, Z. Zeng, and T. Huang, "Adjusting learning rate of memristor-based multilayer neural networks via fuzzy method," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 38, no. 6, pp. 1084–1094, Jun. 2019.
- [21] S. Wen, R. Hu, Y. Yang, Z. Zeng, T. Huang, and Y.-D. Song, "Memristor-based echo state network with online least mean square," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 9, pp. 1787–1796, Sep. 2019.
- [22] Z. Zhang, Y. Zhao, X. Liao, W. Shi, K. Li, Q. Zou, and S. Peng, "Deep learning in omics: A survey and guideline," *Briefings Funct. Genomics*, vol. 18, no. 1, pp. 41–57, 2019.
- [23] Z. Lv, C. Ao, and Q. Zou, "Protein function prediction: From traditional classifier to deep learning," *Proteomics*, vol. 19, no. 14, 2019, Art. no. e1900119.
- [24] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [26] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
- [30] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.
- [31] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1440–1448.
- [32] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.
- [33] K. He, G. Gkioxari, and P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 2961–2969.
- [34] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 779–788.
- [35] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [36] S. Wen, Z. Q. M. Chen, X. Yu, Z. Zeng, and T. Huang, "Fuzzy control for uncertain vehicle active suspension systems via dynamic sliding-mode approach," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 1, pp. 24–32, Jan. 2017.
- [37] S. Wen, T. Huang, X. Yu, Z. Q. M. Chen, and Z. Zeng, "Aperiodic sampled-data sliding-mode control of fuzzy systems with communication delays via the event-triggered method," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 5, pp. 1048–1057, Oct. 2016.
- [38] X. Xie, S. Wen, Z. Zeng, and T. Huang, "Memristor-based circuit implementation of pulse-coupled neural network with dynamical threshold generators," *Neurocomputing*, vol. 284, pp. 10–16, Apr. 2018.
- [39] S. Wen, X. Xie, Z. Yan, T. Huang, and Z. Zeng, "General memristor with applications in multilayer neural networks," *Neural Netw.*, vol. 103, pp. 142–148, Jul. 2018.
- [40] X. Zeng, S. Wen, Z. Zeng, and T. Huang, "Design of memristor-based image convolution calculation in convolutional neural network," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 502–508, 2018.
- [41] S. Xiao, X. Xie, S. Wen, Z. Zeng, T. Huang, and J. Jiang, "GST-memristor-based online learning neural networks," *Neurocomputing*, vol. 272, pp. 677–682, Jan. 2017.
- [42] S. Wang, Y. Cao, T. Huang, and S. Wen, "Passivity and passification of memristive neural networks with leakage term and time-varying delays," *Appl. Math. Comput.*, vol. 361, pp. 294–310, Nov. 2019.
- [43] Y. Cao, Y. Cao, S. Wen, Z. Zeng, and T. Huang, "Passivity analysis of delayed reaction-diffusion memristor-based neural networks," *Neural Netw.*, vol. 109, pp. 159–167, Jan. 2019.
- [44] Y. Cao, S. Wang, Z. Guo, T. Huang, and S. Wen, "Synchronization of memristive neural networks with leakage delay and parameters mismatch via event-triggered control," *Neural Netw.*, vol. 119, pp. 178–189, Nov. 2019.
- [45] S. Wang, Y. Cao, T. Huang, Y. Chen, P. Li, and S. Wen, "Sliding mode control of neural networks via continuous or periodic sampling event-triggering algorithm," *Neural Netw.*, vol. 121, pp. 140–147, Jan. 2020.
- [46] Y. Wang, Y. Cao, Z. Guo, and S. Wen, "Passivity and passification of memristive recurrent neural networks with multi-proportional delays and impulse," *Appl. Math. Comput.*, vol. 0, pp. 1–10, 2019.
- [47] Q. Zou, P. Xing, L. Wei, and B. Liu, "Gene2vec: Gene subsequence embedding for prediction of mammalian N<sup>6</sup>-methyladenosine sites from mRNA," *Proc. RNA*, vol. 25, no. 2, pp. 205–218, 2019.
- [48] L. Wei, Y. Ding, R. Su, J. Tang, and Q. Zou, "Prediction of human protein subcellular localization using deep learning," *J. Parallel Distrib. Comput.*, vol. 117, pp. 212–217, Jul. 2018.
- [49] Z.-K. Feng, W.-J. Niu, and C.-T. Cheng, "Optimizing electrical power production of hydropower system by uniform progressive optimality algorithm based on two-stage search mechanism and uniform design," *J. Cleaner Prod.*, vol. 190, pp. 432–442, Jul. 2018.
- [50] Z. Feng, W. Niu, W. Wang, J. Zhou, and C. Cheng, "A mixed integer linear programming model for unit commitment of thermal plants with peak shaving operation aspect in regional power grid," *Energy*, vol. 12, no. 2179, pp. 1–15, 2019.
- [51] Z. Feng, W. Niu, S. Wang, C. Cheng, and Z. Song, "Mixed integer linear programming model for peak operation of gas-fired generating units with disjoint-prohibited operating zones," *Energies*, vol. 12, no. 1, pp. 1–20, 2019.
- [52] Z. Feng, W. Niu, R. Zhang, S. Wang, and C. Cheng, "Operation rule derivation of hydropower reservoir by k-means clustering method and extreme learning machine based on particle swarm optimization," *J. Hydrol.*, vol. 576, pp. 229–238, Sep. 2019.
- [53] Z. Feng, W. Niu, and C. Cheng, "Optimization of hydropower reservoirs operation balancing generation benefit and ecological requirement with parallel multi-objective genetic algorithm," *Energy*, vol. 153, pp. 706–718, Jun. 2018.
- [54] W. Niu, Z. Feng, M. Zeng, B. Feng, Y. Min, C. Cheng, and J. Zhou, "Forecasting reservoir monthly runoff via ensemble empirical mode decomposition and extreme learning machine optimized by an improved gravitational search algorithm," *Appl. Soft Comput.*, vol. 82, Sep. 2019, Art. no. 105589.
- [55] W. Niu, Z. Feng, C. Cheng, and J. Zhou, "Forecasting daily runoff by extreme learning machine based on quantum-behaved particle swarm optimization," *J. Hydrol. Eng.*, vol. 23, no. 3, pp. 1–12, 2018.

- [56] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 3431–3440.
- [57] G. Wang, M. A. Zuluaga, R. Pratt, M. Aertsen, A. L. David, J. Deprest, T. Vercauteren, and S. Ourselin, "Slic-Seg: Slice-by-slice segmentation propagation of the placenta in fetal MRI using one-plane scribbles and online learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Springer, 2015, pp. 29–37.
- [58] G. N. Stevenson, S. L. Collins, J. Ding, L. Impey, and J. A. Noble, "3-D ultrasound segmentation of the placenta using the random walker algorithm: Reliability and agreement," *Ultrasound Med. Biol.*, vol. 41, no. 12, pp. 3182–3193, 2015.
- [59] G. Wang, M. A. Zuluaga, R. Pratt, M. Aertsen, T. Doel, M. Klusmann, A. L. David, J. Deprest, T. Vercauteren, and S. Ourselin, "Slic-Seg: A minimally interactive segmentation of the placenta from sparse and motion-corrupted fetal MRI in multiple views," *Med. Image Anal.*, vol. 34, pp. 137–147, Dec. 2016.
- [60] A. Alansary, K. Kamnitsas, A. Davidson, R. Khlebnikov, M. Rajchl, C. Malamateniou, M. Rutherford, J. V. Hajnal, B. Glocker, and D. Rueckert, "Fast fully automatic segmentation of the human placenta from motion corrupted mri," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Springer, 2016, pp. 589–597.
- [61] P. Looney, G. N. Stevenson, K. H. Nicolaidis, W. Plasencia, M. Molloholli, S. Natsis, and S. L. Collins, "Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning," in *Proc. IEEE 14th Int. Symp. Biomed. Imag.*, Apr. 2017, pp. 279–282.
- [62] H. Miao, G. Mistelbauer, A. Karimov, A. Alansary, A. Davidson, D. F. Lloyd, M. Damodaram, L. Story, J. Hutter, and J. V. Hajnal, "Placenta maps: In utero placental health assessment of the human fetus," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 6, pp. 1612–1623, Jun. 2017.
- [63] X. Yang, L. Yu, S. Li, X. Wang, N. Wang, J. Qin, D. Ni, and P.-A. Heng, "Towards automatic semantic segmentation in volumetric ultrasound," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Springer, 2017, pp. 711–719.
- [64] G. Wang, M. A. Zuluaga, W. Li, R. Pratt, P. A. Patel, M. Aertsen, T. Doel, A. L. David, J. Deprest, and S. Ourselin, "DeepGeoS: A deep interactive geodesic framework for medical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 7, pp. 1559–1572, Jul. 2019.
- [65] P. Looney, G. N. Stevenson, K. H. Nicolaidis, W. Plasencia, M. Molloholli, S. Natsis, and S. L. Collins, "Fully automated, real-time 3D ultrasound segmentation to estimate first trimester placental volume using deep learning," *JCI Insight*, vol. 3, no. 11, 2018, Art. no. e120178.
- [66] A. G. Howard, M. Zhu, C. Bo, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," Tech. Rep., 2017.
- [67] L. Sifre and S. Mallat, "Rigid-motion scattering for image classification," Ph.D. dissertation, 2014.
- [68] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: <https://arxiv.org/abs/1502.03167>

• • •