Deep Learning-Based Fetal Corpus Callosum Segmentation in Ultrasonic Images

Wen Zheng, Murong Yi, Guiqun Cao, Zhuyu Zhou, and Jian Cheng*

***Abstract*—The corpus callosum is the largest commissural fiber in the cerebral hemisphere that lies at the bottom of the cerebral longitudinal fissure. The Agenesis of Corpus Callosum (ACC) is a congenital disease in the fetal central nervous system malformation, which means the partial or total loss of the corpus callosum during the formation and is detrimental to future development. Its symptom detection mainly depends on the ultrasonic diagnosis, but this method is highly dependent on the experience of doctors because different locations of the fetus and the resolution of the images bring difficulties to the detection of complete callosum. To solve this problem, this paper presents a fusing attention mechanism based on the deep learning method which takes in the advantages of Transformers and dual attention mechanism and realizes accurate semantic segmentation of fetal corpus callosum in ultrasonic images. This method successfully reached an Intersection over Union (IoU) of 59.4%. Besides, this paper also presents the comparison between the performances of different backbone networks and loss functions in order to provide a reference for the application of different parameters according to actual circumstances. Our work provides a reliable reference to locate corpus callosum, thus is promising for the improvement in the diagnosis of ACC and the reduction of the burden of medical workers.**

***Index Terms*—Deep learning, fetal corpus callosum, semantic segmentation, ultrasonic images.**

I. Introduction

The corpus callosum is the largest commissural fiber in the cerebral hemisphere, lying at the bottom of the cerebral longitudinal fissure. The Agenesis of Corpus Callosum (ACC) is a congenital disease in the malformation of the fetal central nervous system. It is the partial or total loss of the fetal corpus callosum during the formation, with an incidence of about 0.03%~0.07%. There is no obvious symptom for this disease during infancy, making it difficult to be detected. However, ACC can lead to future problems like visual disturbance, epilepsy, and dysnoesia [1], which are detrimental to the future development of the fetus and may cast a great burden on the family. As a result, the timely diagnosis of the disease is required for further preparation and treatment.

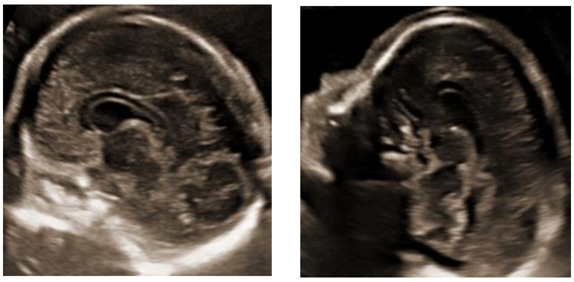
Manuscript received March 3, 2022; revised May 16, 2022.

Wen Zheng and Murong Yi are with the University of Electronic Science and Technology of China, Chengdu, Sichuan, China (e-mail: zwence@163.com, yibayy9@gmail.com).

Guiqun Cao is with Clinical Translational Innovation Center and Molecular Medicine Research Center, West China Hospital of Sichuan University, Chengdu, Sichuan, China (email: caoguiqun@126.com).

Zhuyu Zhou is with Deyang People's Hospital, Deyang, Sichuan, China (e-mail: zhoujin1975@163.com).

*Jian Cheng is the corresponding author and is with the University of Electronic Science and Technology of China, Chengdu, Sichuan, China (e-mail: chengjian@uestc.edu.cn).



(a) (b)

Fig. 1. (a) (left) An ultrasonic image with corpus callosum presented clearly; (b) (right) A disturbed ultrasonic image with a blurry boundary of the corpus callosum.

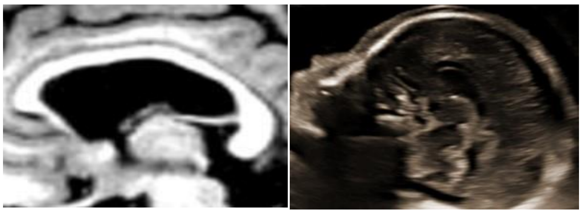
Prenatal ultrasound is acknowledged as one of the most important and useful ways to detect the malformation of the fetal central nervous system. It uses the amplitude of the ultrasonic waves reflected by the placenta and fetus to display the fetus, including the position, size, and shape of its brain. This technology enables the timely diagnosis of ACC, whose value has been gradually realized through clinical practice [2].

A sample ultrasonic image of the fetal corpus callosum is presented in Fig. 1(a). On the sagittal section of the median part, the normal corpus callosum shows a hook-shaped droplet structure; and a backward extension from the top of the septum pellucidum at its end.

The ultrasonic images are a useful basis for diagnosis. However, their quality is easily influenced by liquid and fat, and the images may have poor spatial resolution consequently. Fig. 1(b) is an ultrasonic image whose quality is reduced. Besides, it is more difficult to check the callosum when the fetus is still growing. These factors make the ACC diagnosis a challenging task that greatly relies on the experience of medical workers and may consequently be costive. The application of deep learning methods to assist the segmentation of the fetal corpus callosum can offer a reference for doctors, which may bring a reduction in the workload of medical workers and an increase in the accuracy of ACC diagnosis.

As shown in Fig. 1(b), the disturbed ultrasonic image shows little difference between the corpus callosum and the surrounding tissue, making the corpus callosum hard to be distinguished. Thus, the segmentation of callosum based on ultrasonic images is a challenging task. Many algorithms were raised by former researchers to deal with the problem, roughly divided into *traditional machine learning* algorithms and *deep learning* algorithms. Traditional machine learning algorithms rely on the analysis of the images to raise certain rules for model construction. When the features of the target are well analyzed, effective segmentation can be realized. However, in most cases, analysis of the data is not capable to

DOI: 10.7763/IJCTE.2022.V14.1317



(a) (b)

Fig. 2. (a) (left) A test ultrasonic image used by Ciecholewski;
(b) (right) A disturbed ultrasonic image obtained through clinical practice.

offer enough detailed features to guarantee the model performance, whereas for deep learning algorithms the features are extracted by the model itself through the learning process. With processed input and proper loss function, the networks are able to learn the features and strategy of segmentation themselves, which is more effective and subjective.

In this paper, we will first analyze the up-to-date research progress of traditional machine learning algorithms and deep learning algorithms, then focus on the structure and development of fetal corpus callosum segmentation network. The comparison between the segmentation accuracy of different models will be given as well. Finally, we will analyze the challenges faced in this field and propose the direction of future research.

II. Related Works

A. Traditional Machine Learning Algorithms

Machine learning methods applied for medical image segmentation are mainly active-contour-based segmentation methods [3] and region-based segmentation methods [4].

In 2018, Ciecholewski *et al*. [5] proposed a semi-automatic segmentation method for corpus callosum based on the active-contour-based segmentation method. They compared the performance of three active-contour- based methods on the corpus callosum segmentation, namely: an edge-based active contour model using an inflation/deflation force with a damping coefficient (EM), the Selective Binary and Gaussian Filtering Regularized Level Set (SBGFRLS) method and the Distance Regularized Level Set Evolution (DRLSE) method. However, the above three are all based on traditional semantic segmentation algorithms, which have strict requirements on the setting of initial points and noise. In Fig. 2(a), a corpus callosum image used by Ciecholewski is presented. The resolution of the image is very high, and there is almost no noise. But as shown in Fig. 2(b), an actual ultrasound image of the corpus callosum, low resolution and high noise are challenging factors which make it difficult for the active-contour- based methods to achieve fully automatic segmentation.

In 2021, Qifeng Wang [6] proposed a corpus callosum segmentation method based on a sliding window search algorithm, which used medical prior knowledge to determine the approximate position of the corpus callosum, and then searched the image area through a sliding window to precisely locate the position of the corpus callosum. After determining the position of the corpus callosum, the histogram and bilateral filtering algorithms were used to make the corpus callosum more obvious out of the background, and then the average template map of the corpus callosum was used to finetune the result to obtain the final segmented image. Compared with the work of Ciecholewski, this method can suppress noise to a greater extent and obtain better performance. However, due to the introduction of the sliding window search algorithm and the fitting iterative algorithm, this method is highly time-costing, and is consequently difficult to be applied to the actual production environment.

In summary, the traditional algorithm is very dependent on the setting of the initial value, with weak ability to resist noise. It is also very dependent on the experience of scientific researchers. The models have low accuracy and low model generalization. Therefore, in this paper we choose to use deep learning-based segmentation method as the main solution.

B. Deep Learning Algorithms

Currently, deep learning methods applied for medical image segmentation are mainly divided into 2 categories. One is the *Encoder-Decoder* structure, and a typical network is the Unet [7]. Another is the *backbone network* structure, and a typical network is the DeepLab [8].

The Unet has been widely used in Biomedical Data Science ever since its release in 2015. Its popularity can be attributed to its characteristics: the *Encoder-Decoder* structure and the *jump-connection*. The encoder of the network down-samples the input image 4 times, reducing the resolution to one-sixteenth of its original size; accordingly, the decoder up-samples the high-level semantic feature map 4 times, recovering the resolution to the initial. Compared with Fully Convolutional Network (FCN) and other networks applied for semantic segmentation, Unet innovatively operates jump-connection in each layer, fusing the features of different scales. This method enables a more accurate segmentation.

DeepLab was first released by Liang-Chieh Chen, George Papandreow, Florian Schroff, and Hartwig Adam. Then adapted versions, DeepLabV2 and V3 were raised [9], [10]. DeepLab embeds context information of different scales to improve the consistency between the net and the spatial pyramid pooling module. The early versions of DeepLab did not take the jump connection, which resulted in the loss in detail information. Therefore, DeepLabV3+ [11] referred to Unet to solve this problem by involving jump connection.

Transformers [12] was first proposed by Google for machine translation, which had very good results for Natural Language Processing (NLP). Later, researchers found that modified Transformers can also be used for vision tasks. TransUNet [13] proposed by Jieneng Chen *et al*. in 2021 is the first network frame for medical image segmentation that takes advantage of Transformers. On the one hand, the feature blocks output by the feature extraction network are encoded as the input sequence of Transformer for features. On the other hand, it realizes semantic segmentation of medical images by referring to the structure of Unet: the decoder up-samples the encoded features and fuses them with the high-resolution feature maps.

In conclusion, U-shaped structure and jumping connection are both applied in DeeplabV3 and TransUnet, with different designs to obtain more reception field and multi-scale information. The performance of different networks on corpus callosum segmentation will be compared, and another updated method will also be presented in this paper.

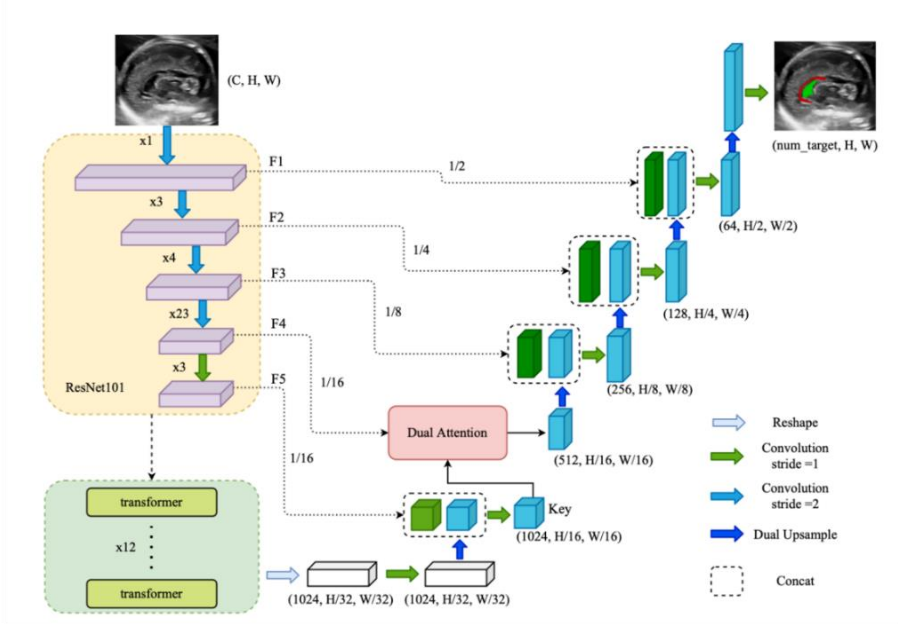


Fig. 3. The structure of the proposed network.

III. Method

A. Model Design

The function of our network is to output an accurate segmentation map of the input ultrasonic image. The input, ultrasonic images of the fetal head, have the resolution of $H \times W$ and channels of C, which is marked as $x \in R^{H\times W\times C}$. The output of the network is a map with the resolution of $H \times W$.

Our network combines Transformers and Convolutional Neural Network (CNN) to extract features. It is an idea inspired by TransUnet. The CNN extractor is followed by 12 layers of Transformers to balance the local information and global information. Considering the images are greatly disturbed by noise, we took ResNet101 [14] as the CNN feature extractor since its deeper structure enables stronger ability as an extractor. Inspired by dual attention mechanism [15], the decoder takes the output features of the transformers as the key values of the dual attention to fuse spatial features and channel features. This fusion brings great improvements to the network performance. Besides the segmentation of the corpus callosum, we also take the transparent compartment, a tissue next to the callosum, as another target because it is also an important basis for the diagnosis of ACC. The complete structure of the network is shown in Fig. 3.

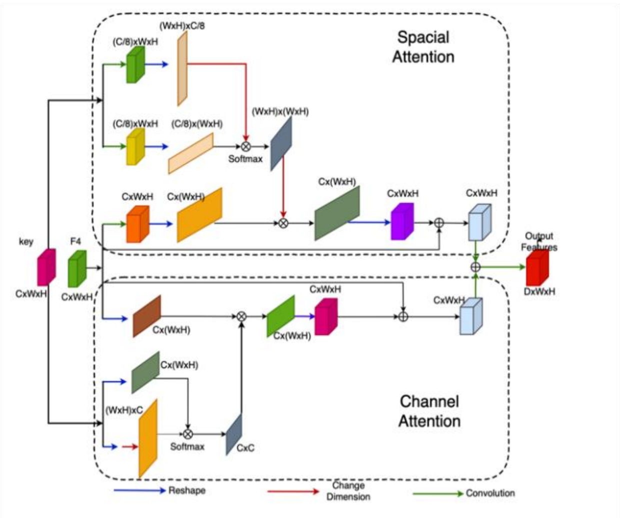


Fig. 4. The structure of the dual attention block.

For the first 3 layers of ResNet, convolution kernels with a stride of 2 and a size of 3 × 3 are used to achieve down-sampling and extract high-level features. Inspired by dilated convolution [16], at the last layer of ResNet, dilated kernels of 3 × 3 size and stride of 1 are applied to enlarge the reception field, which is conducive to the extraction of more information. As illustrated in Fig. 3, the dual attention block takes $F4$ and key as the input, and output D channels of feature maps with the same resolution. The key is convolved and resized before being taken as the weight values of spatial attention and channel attention, then multiply with the preprocessed input features. The final output results of the two channels are concatenated and processed by a CNN network. The detailed process is shown in Fig. 4.

B. Loss Function

The loss function is a critical strategy that influences the performance of network training. For our task, the segmentation of fetal corpus callosum, the quantity of the background and target pixels are greatly imbalanced, so the flexible adjustment to the weight of False Positive (FP) and False Negative (FN) predictions is needed. As a result, *Tversky Loss* [17] is taken as the loss function. Its expression is presented in (1).

$$T(A,B) = \frac{A \cap B}{A \cap B + \alpha|A-B| + \beta|B-A|} \quad (1)$$

where A and B stands for the prediction and ground truth respectively; $|A-B|$ is the false positive and $|B-A|$ is the false negative. By finetuning α and β, we can adjust the penalties of FP and FN to the training loss, which affects the final performance. In our experiment, α is set as 0.6 while β is set as 0.4.

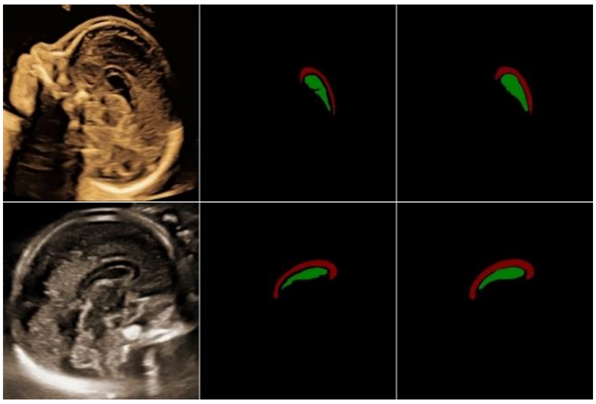


Fig. 5. The input images (left), ground truths (middle), and predictions (right).
The callosum is marked in red and the transparent compartment is marked in green.

IV. Experiment and Discussion

A. Dataset and Experiment Setup

The dataset is obtained from People's Hospital of Deyang City and has 443 samples in total. The ultrasonic images were manually labeled, with the fetal brain, corpus callosum, and transparent compartment being marked. The ultrasonic images were firstly cropped into input imageswith the resolution of 512 × 512. Due to the limited data, image augmentation was applied, including random sharpening, histogram equalization, flipping, and rotation. The network was trained on a GTX TITAN X, optimized by SGD optimizer. The learning rate and batch size were set as 0.004 and 4 respectively, and the network was trained for 200 epochs.

B. Experiments and Discussion

Sample outputs of our model is shown in Fig. 5. The input ultrasonic images are on the left; the ground truth labels are in the middle and the predictions (the output of our network) are on the right. It can be demonstrated that our model can realize the segmentation of fetal corpus callosum very well.

Further observation of the result illustrates that the predicted corpus callosum area is slightly larger than the ground truth. This is caused by the greater shape sensitivity than boundary sensitivity in this task because of the fine-tuned parameters in the loss function. The reason behind the greater shape sensitivity is that we need to prevent the segmentation from being discontinuous due to the thinness of the corpus callosum, which is corresponding to the expected results. The adjustment on parameters applied to the Tversky Loss was to balance the penalty of FP and FN to make the model more inclined to ensure the correctness of the predicted shape.

To have a better demonstration of the performance of our model on the dataset, we cast experiments on the same dataset using 3 typical networks which were once the state-of-the-art (SOTA) models: DeepLabV3, Unet, and TransUnet. We will also compare the different model performances when ResNet50 and Dice Loss are applied.

As shown in Table I, the feature extractor and the loss function of DeepLabV3, TransUnet, and our model are all ResNet101 and Tversky Loss. The Intersection over Union (IoU) index is used to measure the performance of the models.

TABLE I: Performances of Different Networks on the Dataset (CC Stands for Corpus Callosum and TC Stands for Transparent Compartment)

Network	Mean IoU	IoU of TC	IoU of CC
DeepLabV3	63.7%	69.3%	58.1%
Unet	56.1%	61.9%	50.3%
TransUnet	63.5%	69.4%	57.6%
Our Model	**64.5%**	**69.5%**	**59.4%**

TABLE II: Different Performances of Feature Extractors

Network	Mean IoU	IoU of TC	IoU of CC
DeepLabV3+ResNet50	61.7%	67.6%	55.8%
DeepLabV3+ResNet101	63.7%	69.3%	58.1%
Our Model+ResNet50	62.6%	66.7%	59.0%
Our Model+ResNet101	**64.5%**	**69.5%**	**59.4%**

TABLE III: Different Performances of Loss Functions

Network	Mean IoU	IoU of TC	IoU of CC
DeepLabV3+Dice Loss	59.9%	66.7%	53.1%
DeepLabV3+Tversky Loss	63.7%	69.3%	58.1%
Our Model+Dice Loss	61.7%	67.4%	55.9%
Our Model+Tversky Loss	**64.5%**	**69.5%**	**59.4%**

It is shown in the results that our model has the best performance. It does not bring great improvement in the segmentation of transparent compartment compared with DeepLabV3 and TransUnet. However, it enjoys an increase of 1%~2% approximately in the segmentation of the corpus callosum and about 1% in the mean Intersection over Union (mIoU) index.

In Table II, we present the results of using different feature extractors (ResNet50 and ResNet 101) for DeepLabV3 and our network, while the loss functions are all Tversky Loss. The results show that compared with ResNet50, ResNet101 brings an increase of about 2%~3% in the segmentation of transparent compartment, and of about 2% in the segmentation of the corpus callosum. Concluded from Table II, ResNet101 can improve the performance by about 2% compared with ResNet 50.

In Table III, ResNet101 is fixed as the feature extractor for DeepLabV3 and our model, to compare the different performance of Dice Loss ($\alpha = 0.5,\ \beta = 0.5$) and Tversky Loss ($\alpha = 0.6,\ \beta = 0.4$). From the table, Tversky Loss has an obvious advantage over the Dice Loss, with a 2%~3% improvement for transparent compartment segmentation and a 4%~5% improvement for corpus callosum segmentation. The better performance of Tversky Loss can be attributed to its resistance to imbalanced samples brought by fine-tuned parameters.

So far as we have analyzed, we can conclude that for the corpus callosum segmentation task, our network outperforms DeeplabV3 and TransUnet by 1.3% and 1.6% in corpus callosum detection respectively. We can also conclude that the Tversky Loss outperforms the Dice Loss, and Resnet101 outperforms Resnet50. However, ResNet101 has more parameters. Taking this into account, feature extractors should be chosen according to actual needs.

V. Conclusion

This paper proposed a segmentation algorithm for fetal corpus callosum based on deep learning, which combines the advantages of Transformer and dual attention mechanism to improve the performance of the network. Our model reached 59.4% of IOU on the corpus callosum segmentation task, which is promising for future application. The results of ultrasound segmentation of the fetal corpus callosum can provide doctors with a valuable reference, which is conducive to reducing the workload of doctors and improving the accuracy of the diagnosis of ACC.

To realize the end-to-end diagnosis in the assistance of the work in this paper, two directions of future work are given: 1. Improve the performance of the model by making full use of the unlabeled dataset based on self-supervised learning; 2. On the basis of the results of automatic segmentation of corpus callosum, raise criteria for the diagnosis of ACC using statistical method. Meanwhile, we will also make efforts to collect more data in order to introduce more automation in the detection of ACC.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

All authors had approved the final version.

Acknowledgment

This research was supported by Sichuan Science and Technology Program (No. 2020YFG0085, 2021YFG0328), and in part by the National Natural Science Foundation of China (No. 62071104).

References

[1] X. Li and Q. Wang, "Magnetic Resonance Imaging (MRI) diagnosis of fetal corpus callosum abnormalities and follow-up analysis," *J. Child Neurol.*, vol. 36, no. 11, pp. 1017-1026, Oct. 2021.

[2] K. Ono, Y. Iwamoto, Y.-W. Chen, and M. Nonaka, "Automatic segmentation of infant brain ventricles with hydrocephalus in MRI based on 2.5D u-net and transfer learning," *Journal of Image and Graphics*, vol. 8, no. 2, pp. 42-46, June 2020.

[3] J. Dong *et al.*, "Local-global active contour model based on tensor-based representation for 3D ultrasound vessel segmentation," *Physics in Medicine & Biology*, vol. 66, no. 11, 2021

[4] L. Zou *et al.*, "A survey on regional level set image segmentation models based on the energy functional similarity measure," *Neurocomputing*, 2020.

[5] M. Ciecholewski and J. H. Spodnik, "Semi–automatic corpus callosum segmentation and 3d visualization using active contour methods," *Symmetry*, vol. 10, no. 11, 2018

[6] Q. Wang, "Automatic extraction of fetal corpus callosum and cerebellar vermis ultrasound images," Ph.D. dissertation, Dalian University of Technology, 2021.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Springer International Publishing*, 2015.

[8] N. Shigeta, M. Kamata, and M. Kikuchi, "Effectiveness of pseudo 3D feature learning for spinal segmentation by CNN with u-net architecture," *Journal of Image and Graphics*, vol. 7, no. 3, pp. 107-111, September 2019.

[9] L. C. Chen, G. Papandreou, I. Kokkinos *et al.*, "DeepLab: Semantic image segmentation with deep convolutional nets," *Atrous Convolution, and Fully Connected CRFs*, 2016.

[10] L. C. Chen, G. Papandreou, F. Schroff, *et al.*, "Rethinking atrous convolution for semantic image segmentation," *Computer Vision and Pattern Recognition*, 2017.

[11] L. C. Chen, Y. Zhu, G. Papandreou, *et al.*, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. European Conference on Computer Vision*, 2018, pp. 801-818.

[12] A. Vaswani, N. Shazeer, N. Parmar, *et al*., "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[13] J. Chen, Y. Lu, Q. Yu, *et al*., "Transunet: Transformers make strong encoders for medical image segmentation," *Computer Vision and Pattern Recognition*, 2021.

[14] K. He, X. Zhang, S. Ren, *et al.*, "Deep residual learning for image recognition," in *Proc. the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778.

[15] J. Fu, J. Liu, H. Tian, *et al.*, "Dual attention network for scene segmentation," in *Proc. the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3146-3154.

[16] P. Wang, P. Chen, *et al*., "Understanding convolution for semantic segmentation," in *Proc. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018, pp. 1451-146.

[17] S. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *Proc. International Workshop on Machine Learning in Medical Imaging*, 2017.

Wen Zheng was born on November 4, 1995, is currently a graduate student at the School of Information and Communication Engineering, University of Electronic Science and Technology of China.

He majors in electronics and communication engineering and will receive his master's degree in January 2022. His research interest is computer vision.

Murong Yi was born on 24th August in 2000, is currently an undergraduate student at the University of Electronic Science and Technology of China, majoring in communication engineering, and will receive her scholar degree in June 2022.

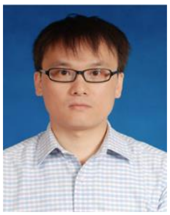
Her research interests include machine learning, computer vision, medical image analysis, and video analysis.

Guiqun Cao received the M.Bio and Ph.D degree in 2012 and 2017 from the Zunyi Medical University and West China School of Medicine, Sichuan University respectively. Currently, she is working at West China Hospital of Sichuan University.

Her main research interests include biochemistry molecular biology and medical artificial intelligence, with a focus on AI that simulates human doctors for disease diagnosis.

Zhuyu Zhou received her master's degree in imaging medicine and nuclear medicine from Shanghai Jiao Tong University in 2012. From 2013 to 2020, she served as deputy chief physician in People's Hospital of Deyang City, Sichuan, China. She is currently the chief physician of People's Hospital of Deyang City.

Her research interests include prenatal ultrasound diagnosis, fetal congenital heart disease etc.



Jian Cheng received his Ph.D. degree in pattern recognition and intelligent system from Shanghai Jiao Tong University, Shanghai, China, in 2006.

From 2006 to 2007, he was an assistant researcher with the Chengdu Information Technology of Chinese Academy of Sciences Co., Ltd. He is currently a professor at the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China.

His research interests include machine learning, computer vision, remote sensing image analysis, medical image analysis, multimodal image classification, video surveillance and scene understanding, human behavior analysis, etc.